0787

PB 204 180

COMPUTER MAPPING
as an aid in
AIR POLLUTION STUDIES

VOLUME TWO: INDIVIDUAL REPORTS

REPORT L
SYMAP INTERPOLATION CHARACTERISTICS

by

Donald S. Shepard

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

A

70

BIBLIOGRAPHIC DATA SHEET	1. Report No. APTD-0787	2.	3. Recipient's Accession No.
4. Title and Subtitle Computer Mapping as an aid in Air Pollution Studies Volume Two: Individual Reports - Report L - Symap Interpolation Characteristics			5. Report Date April 1970 6.
7. Author(s) Donald S. Shepard			8. Performing Organization Rept. No.
9. Performing Organization Name and Address Graduate School of Design Harvard University Cambridge, Massachusetts 02138			10. Project/Task/Work Unit No. 11. Contract/Grant No. 68A-2405D
12. Sponsoring Organization Name and Address National Air Pollution Control Administration U. S. Public Health Service U. S. Department of Health, Education, and Welfare Washington, D. C. 20201			13. Type of Report & Period Covered 14.

15. Supplementary Notes DISCLAIMER - This report was furnished to the Office of Air Programs by Graduate School of Design, Harvard University, Cambridge, Massachusetts 02138 in fulfillment of Grant No. 68A-2405D

16. Abstracts
A technical discussion of Symap interpolation characteristics is presented in four sections; The Symap interpolation algorithm, Sensitivity and accuracy of interpolation, Use of electives to affect interpolation, and Use of barriers to affect interpolation. Appendices on Symap compared to other methods of interpolation and on a map comparison program are included.

17. Key Words and Document Analysis. 17a. Descriptors
Air pollution
Computer programs
Mapping
Algorithms
Mathematical logic

17b. Identifiers/Open-Ended Terms
Symap

17c. COSATI Field/Group 13B

18. Availability Statement Unlimited	19. Security Class (This Report) UNCLASSIFIED	21. No. of Pages 71
A	20. Security Class (This Page) UNCLASSIFIED	22. Price

LABORATORY FOR COMPUTER GRAPHICS
AND SPATIAL ANALYSIS

Graduate School of Design
Harvard University
Cambridge, Massachusetts

This report was funded by:

Demonstration Grant 68A-2405D
National Air Pollution Control Administration
U. S. Public Health Service
U. S. Department of Health, Education, and Welfare

John C. Goodrich, Editor
Gail Howrigan, Associate Editor

Rosalie Alexander
Carol Chase
Candace Howes
Whitmore John
Linda Peacock

April 1970

COMPUTER MAPPING

as an aid in

AIR POLLUTION STUDIES

Personnel

Principal Investigators

Howard T. Fisher	Project Director (Associate Director of The Laboratory for Computer Graphics and Spatial Analysis)
John C. Goodrich	Associate Project Director
William Warntz	(Director of The Laboratory for Computer Graphics and Spatial Analysis)
Allan H. Schmidt	(Associate Director of The Laboratory for Computer Graphics and Spatial Analysis)

Case Study Investigators

Norma Bogen	St. Louis Project
Paul Jack Dangermond	California and Puget Sound Studies
Steven Fosburg	Statistical Investigation
Jean Maurice Granger	Montreal Project
Donald Shepard	St. Louis Project/Program Investigation
Elia Werczberger	Kansas City Project

Programming Support and Development

Chang-I Hua
Lannon Leiman
Kathleen Reine

Research Assistants

Whitmore John
Johnny Kirkland
Helen Mansfield
Adrian Thomas

Ellen Bussard
Robert Cartwright
Katherine Mickey
Warnie Webster

Administrative and Secretarial Support

Malcolm Fitzpatrick
Gail Howrigan
Linda Peacock
Laurence Yont

Regina McDonald
Donna L. Moorhead
Claudia S. Rosen
Judith Strong

PREFACE

From July 1, 1968 through December 31, 1969, the Laboratory for Computer Graphics and Spatial Analysis conducted a demonstration of several conceptual, methodological, and technical approaches to problems in the field of air pollution, employing computer-produced maps and allied displays. This work was done under a grant (No. 68A-2405D) from the Division of Control Agency Development, National Air Pollution Control Administration (NAPCA), Public Health Service, U. S. Department of Health, Education, and Welfare.

The results of this study, taken in its entirety, should make an important contribution to the work of Federal, state, and local air pollution agencies, especially in their efforts to determine the best control strategies for use in air quality control regions. The computer is an ideal tool for relating and combining sets of data, and for considering alternatives. For instance, graphic displays can contribute significantly to the evaluation of both short-term and long-term control strategies.

In demonstrating "the use of computer graphics as a means for more clearly portraying data related to the sources, distribution, and community effects of air pollution," the primary emphasis was on case studies for specific areas. The secondary emphasis, to support and complement the case studies, included:

(a) "the development of instructional manuals covering the procedures required for the preparation and interpretation of computer maps of various kinds, as related to air pollution,"

(b) "the use of certain newly developed techniques for describing, analyzing, and portraying quantitative surfaces and movements relative to such surfaces, as related to air pollution," and

(c) "an appraisal of alternative equipment, techniques, and computer programs presently or potentially available for the production of computer graphic displays pertinent to air pollution."

The case study regions were selected in consultation with NAPCA. Considerations included:

(a) sufficient air quality, emission, and related socio-economic and meteorological data--spatially distributed over the region--to enhance the likelihood of adequately demonstrating the mapping techniques;

(b) representative problems in data handling and mapping so that the techniques demonstrated would have applicability to as many other areas as possible;

(c) special problems in data handling and mapping so that as many techniques as possible could be adapted, developed, and tested;

(d) areas of which NAPCA and the individual investigators had special knowledge and could therefore utilize relevant research to enhance the study;

(e) areas where an immediate interest in mapping had been expressed for on-going Federal, state, and local air pollution activities.

Other research was undertaken to support and complement the case studies, and, in particular, to provide:

(a) the widest possible distribution of the mapping programs and accompanying explanation of the techniques through

i) copies of the Laboratory's SYMAP Computer mapping program,
ii) an introductory correspondence course in the SYMAP program,
iii) general program description and supplementary explanation for the existing user-manuals;

(b) a general review of present and potential hardware, output devices, and programs applicable to computer graphic displays of air pollution and related data;

(c) means for predicting, projecting, and correlating air pollution and related data, utilizing the latest

knowledge in planning, geography, and spatial analysis in general;

(d) dissemination of information on approaches and techniques to Federal, state, and local air pollution agencies, especially for use in the air quality control regions.

The final report is divided into two parts: Volume One: Summary, and Volume Two: Individual Reports--the detailed reports on the case study areas and the technical reports related to other research undertaken. The pertinent graphic displays are included with the appropriate individual report and a complete set is provided with the Summary Volume.

Introductory sections describing the air pollution and related data used are provided in the major case study reports for St. Louis and Kansas City; these should be referred to by anyone unfamiliar with the types of data and problems existing in this field. Detailed discussions of special data handling and mapping techniques are presented either as appendices to case study reports, or in separate technical reports.

The Summary Volume includes a discussion of all the areas of research undertaken and contains the complete table of contents for the reports of Volume Two. This should provide a useful reference of the special problems and technical discussions of one report which may be of interest to the reader of another detailed report.

The technical discussion of Symap interpolation characteristics is presented here as Report L of Volume Two. The report is divided into four sections -- The Symap interpolation algorithm, sensitivity and accuracy of interpolation, use of electives to affect interpolation, and use of barriers to affect interpolation. Appendices on Symap compared to other methods of interpolation and on a map comparison program are included.

The various studies discussed in the report were carried out principally by Donald Shepard, with the assistance of Katherine Mickey. The aid of Thomas K. Peucker, Assistant Professor at Simon Fraser University, Burnaby, British Columbia, and William Warntz of the Laboratory, is gratefully acknowledged for the work discussed in section two.

The report was parepared principally by Donald Shepard.

John C. Goodrich, Editor
Cambridge, Massachusetts
April, 1970

REPORT L
SYMAP INTERPOLATION CHARACTERISTICS

Table of Contents

1. THE SYMAP INTERPOLATION ALGORITHM

1.1 Computation Model
1.2 Modifications
1.2.1. Search Radius
1.2.2. Direction
1.2.3. Slopes

1.3. Program Organization and Characteristics
1.3.1. Barriers
1.3.2. Sample Computation Using Algorithm
1.3.3. Summary and Evaluation

2. SENSITIVITY AND ACCURACY OF INTERPOLATION

2.1 Sensitivity to Total Number of Data Points
2.2. Distribution of Error
2.3. Improvements in the Interpolation Algorithm
2.3.1. Fixed Number of Data Points
2.3.2. Inverse Distance Function
2.3.3. Comparison with Standard Procedure
2.3.4. Conclusions

3. USE OF ELECTIVES TO AFFECT INTERPOLATION

3.1. Properties of SYMAP interpolation
3.2. Value Extremes
3.2.1. Plateau Effect
3.2.2. Fractional Extrapolation
3.3. Search Radius
3.3.1. Initial Radius
3.3.2. Maximum Radius
3.4. Detail of Surface Computation
3.4.1. Number of Data Points
3.4.2. Independent Computation

4. USE OF BARRIERS TO AFFECT INTERPOLATION
4.1 Crossing Barriers
4.2 Impermeable Barriers
4.2.1 Dividing Line
4.2.2. Closed Loop

4.3. Permeable Barriers
4.3.1. Dividing Line
4.3.2. Protruding Barrier
4.3.3. Other Configurations

REFERENCES

APPENDIX 1 SYMAP COMPARED TO OTHER METHODS OF INTERPOLATION

APPENDIX 2 MAP COMPARISON PROGRAM

List of Illustrations

Section 1. THE SYMAP INTERPOLATION ALGORITHM

Example 1-1	Sample Computation Using Interpolation Algorithm

Section 2. SENSITIVITY AND ACCURACY OF INTERPOLATION

Table 2-1	Sensitivity and Accuracy of Interpolation:
Example 2-1	Old Interpolation
Example 2-2	Frequency Distribution of Differences
Example 2-3	Interpolation with 4 Points
Example 2-4	Theoretical Surface Grid One
Example 2-5	Theoretical Surface Grid One A
Example 2-6	Theoretical Surface Grid Two
Example 2-7	Theoretical Surface Grid Three
Example 2-8	Theoretical Surface Grid Four
Example 2-9	Map 2--4 Data Points (New Interpolation)
Example 2-10	Map 5--272 Data Points (New Interpolation)

Section 3. USE OF ELECTIVES TO AFFECT INTERPOLATION

Table 3-1	Co-Linear Data Points
Example 3-1	Interpolation Algorithm 3 co-linear Data Points
Example 3-2	Interpolation Aglorithm 3 Co-linear Data Points oblique view of surface
Example 3-3	Interpolation Algorithm 3 Co-linear Data Points Profile
Example 3-4	Interpolation Algorithm 3 Co-linear Data Points Profile Fractional Extrapolation 0.0
Example 3-5	Interpolation Algorithm 3 Co-linear Data Points Profile Fractional Extrapolation 0.2

Section 4. USE OF BARRIERS TO AFFECT INTERPOLATION

Example 4-1	Crossing Barriers Odd number of Times
Example 4-2	Crossing Barriers Even Number of Times

Example 4-3 Permeable Barrier Protruding Type
Example 4-4 Impermeable Barrier Dividing Line Type Strength -1
Example 4-5 Permeable Barrier Dividing Line Type Strength 1
Example 4-6 Permeable Barrier Dividing Line Type Strength 5

1. THE SYMAP INTERPOLATION ALGORITHM

The potential user of SYMAP will best be able to understand the relation of the SYMAP algorithm to other methods of interpolation through a general understanding of the assumptions and methods of the program. The FORTRAN IV program was developed independently of any particular field to which it might be applied and consequently produces a surface with properties intuitively desirable in the absence of detailed information.[1]

1.1 COMPUTATION MODEL

The computation part of the program takes as given the coordinates of one or more "data points" and the values associated with them. Unless otherwise specified, the surface is smooth (continuously differentiable), passes through the given points, and represents trends or patterns which those points imply. The method of calculation is a weighted average of slopes and values of nearby data points, developed from a gravity-type model.

According to the basic model, the value at a point P would be a weighted average of the values at data points 1, 2,....n, with the weighting based on the inverse square[2] of the distance.

Let $\overline{P1}$ be the distance from P to data point 1
$\overline{P2}$ be the distance from P to data point 2
Z_1 be the value at data point 1
Z_2 be the value at data point 2
Z_p be the computed value at the point P.

$$\text{Then } Z_p = \frac{\frac{1}{(\overline{P1})^2} Z_1 + \frac{1}{(\overline{P2})^2} Z_2 + \frac{1}{(\overline{P3})^2} Z_3 + \cdots + \frac{1}{(\overline{Pn})^2} Z_n}{\frac{1}{(\overline{P1})^2} + \frac{1}{(\overline{P2})^2} + \frac{1}{(\overline{P3})^2} + \cdots + \frac{1}{(\overline{Pn})^2}}$$

$$\text{or } Z_p = \frac{\sum_{i=1}^{n} \frac{1}{(\overline{Pi})^2} Z_i}{\sum_{i=1}^{n} \frac{1}{(\overline{Pi})^2}}$$

If the point P is very near data point 1, for example, then $\overline{P1}$ is small compared to $\overline{P2}$ and $\overline{P3}$, etc., so that the weighting $\frac{1}{\overline{P1}^2}$ is large compared to $\frac{1}{\overline{P2}^2}$ and $\frac{1}{\overline{P3}^2}$, etc.

Therefore the value at P is the limit of

$$\frac{\frac{1}{(\overline{P1})^2} Z_1}{\frac{1}{(\overline{P1})^2}} = Z_1$$

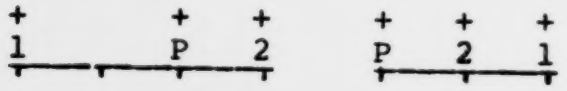
as the distance $\overline{P1}$ tends to zero, so values near the data point 1 are near the value Z_1.

It can be shown that by computing such a value Z_p for every point P in the x-y plane, one obtains a "smooth" surface. However, this method contains the following shortcomings:

(i) If the number of data points is large, the calculation of Z becomes correspondingly longer; the method becomes inefficient or impractical.

(ii) Direction was not considered. The value obtained at P -- depending only on the distances to the points 1, 2, 3,....n -- would

be the same regardless of whether the points were spaced about P, or all lay roughly in the same direction. With direction not considered, the following configurations of data points would give identical values at P, assuming $\overline{P1}$ and $\overline{P2}$ remain unchanged:



(iii) Because of the particular weightings employed, the surface is level at every data point, regardless of the location of or value at that point.[3]

As the method of weighted averages was well suited to computer application and satisfied the basic criteria, it was retained as the basis for the SYMAP program, with several important modifications.

1.2 MODIFICATIONS

1.2.1. Search Radius

The amount of computation was limited by computing Z_p from only the "nearby" data points. Based on the number of data points and the area over which they extend, an initial search radius, R, is determined,[4] such that a circle of that radius will generally contain the average number of points on which interpolation is to be based.[5] At any character location P, those data points whose effective distance from P is less than R are first selected. If more than the allowed

maximum number of points are found, the furthest are eliminated by contracting the search radius until only the allowed number remain. If fewer than the allowed minimum number of data points are found within the initial search radius, the search radius, R, is expanded until the specified minimum number of points is reached, the user-specified maximum search radius is reached, or until all points not blocked by impermeable barriers are used.(6) The inverse distance weightings are used for data points near P. At greater distances, the weighting is specially determined -- as the distance to a data point approaches the final search radius, R', the weighting of that point tends to zero (0). One may think of data points beyond the search radius as also having a weighting of 0; eliminating points outside the search radius causes no discontinuity in the weightings or in the surface produced.

1.2.2 Direction

So that the relative locations of data points enter the computation -- in addition to relative distances -- a "directional isolation" of data point "i" is found by the sum(7)

$$Q_i = \frac{1}{p1}\{1-\cos(iP1)\} + \frac{1}{p2}\{1-\cos(iP2)\} + \ldots + \frac{1}{pn}\{1-\cos(iPn)\}$$

If the other data points -- 1, 2, 3 -- lie in about the same direction from P as "i" does, then the angles iP1, iP2, iP3 are small; the quantities 1-cos(iP1) are also small, and the sum Q_i is also close to zero. Since "i" is not the only point in a particular direction, it is given a reduced weighting, defined as follows:

Let the distance from location P to data point j be $\overline{Pj}$. Then the "distance weighting is (7)

$$S_j = \begin{cases} \dfrac{1}{\overline{Pj}} & \text{for } 0 < \overline{Pj} \leq R'/3 \\ f(\overline{Pj}) & \text{for } R'/3 < \overline{Pj} \leq R' \end{cases}$$

Let $H = \sum_j S_j$ and let $T_i = \sum_{j \neq i} S_j \times \{1-\cos(iPj)\}$, where the points j are neighboring data points within the search radius R'. Then the total weighting at location P of data point i is:

$$W_i = (S_i)^2 \times \{H + T_i\}$$

The greater a data point's "directional isolation" and the smaller its distance, the greater its weighting and hence its influence in determining the interpolated value.

1.2.3 Slopes

To prevent the occurrence of a level surface at data points, a "slope" -- or two-dimensional gradient $\frac{\partial z}{\partial x}\hat{i} + \frac{\partial z}{\partial y}\hat{j}$ -- is estimated at each data point by taking a weighted average of the slopes of several secant planes. Each plane contains the point P and one of its neighboring data points i; the plane is as nearly horizontal as possible -- the line from P to the neighboring data point would be the line of steepest slope in that plane. The value of the surface at the point P approaches $z'_i = Z_i + \Delta Z_i$. Here $\Delta Z_i = \{\frac{\partial z}{\partial x}\Big|_i \Delta x + \frac{\partial z}{\partial y}\Big|_y \Delta y\} \times K_i$

where Δx and Δy are the x and y distances, respectively, from i. The factor, $K_i = \frac{a}{a + \overline{Pi}}$, is inserted so that the effect of considering the slope is kept small at large distances. The parameter a is chosen so that even if i were the data point with the steepest slope, ΔZ would still be less in magnitude than a specified fraction of the total range of Z.[8]

If a data point i is higher than the surrounding data points, the slope at the point will not be zero unless the weighted slopes of the secant planes exactly cancel. Thus the surface will continue to rise in some direction. As a consequence, relative maxima and minima often do not occur at data points. Unless requested otherwise by the user, the program separately restricts interpolated values to the range of the input data values.

The interpolation algorithm gives characteristic effects when extrapolating beyond the study area. As the distance increases, all the nearest data points have roughly the same distance and direction, resulting in approximately equal weightings. Thus, the values will tend towards a limit near the average of the closest data points.[9]

1.3 PROGRAM ORGANIZATION AND CHARACTERISTICS

For efficiency, the foregoing procedure is normally used to calculate the values only at grid points -- every third character location across and every other down.[10] The intermediate points are evaluated by a

quick linear interpolation between the grid points.[11] For accuracy, all values are computed as decimal (floating point) numbers and converted to levels and symbols just before printing.

The value at a character or grid location containing a data point is the modified value, Z', for that data point. A tolerance of 1 1/2 character locations is used in assigning data points to grid points. If several data points fall at a character location, the value used is their average.[12]

1.3.1 Barriers

If barriers to interpolation are included, the algorithm is modified. The distances used in calculating weightings are "effective distances." The square of the effective distance equals the square of the actual distance, plus the sum of the squares of strengths of all barriers crossed an odd number of times. Impermeable barriers are assigned a large strength so that the effective distance to a data point across an impermeable barrier is so large that the point is not considered a nearby data point. Interpolation from data points across the impermeable barrier is thus eliminated.[13]

1.3.2 Sample Computation Using Algorithm

As shown in Example 1-1, suppose that the population density is 200 people/square mile around data point 1, and 400 people/square mile around data point 2. The SYMAP program would estimate the population density

at P, between {1} and {2}, as follows:

$$Z_p = \frac{Z_1 W_1 + Z_2 W_2}{\Sigma W_i} \quad , \text{ or}$$

$$Z_p = \frac{(217.4)(.875) + (384.6)(2.5)}{3.375} = 341$$

If one used linear interpolation between the points, the value would be 333. Although the value obtained by the SYMAP algorithm is almost the same, the latter is applicable to a wide variety of configurations.

EXAMPLE 1-1

SAMPLE COMPUTATION USING INTERPOLATION ALGORITHM

i	S_i	H	T_i	W_i
{1}	.5	1.5	2.0	.875
{2}	1.0	1.5	1.0	2.5

i	Z_i	$\frac{\partial z}{\partial x}\|_i$	K_i	ΔZ_i	Z_i
1	200	66.7	.261	17.4	217.4
2	400	66.7	-.231	-15.4	384.6

$\Sigma w_i = 3.375$

1.3.3 Summary and Evaluation

The value at a normal grid point, P, of the surface is

$$Z_p = \frac{\sum_i w_i Z'_i}{\sum_i w_i}$$

where w_i is the weighting of the data point i, Z'_i is the value at data point i -- modified by the slope at i and its displacement from point P. This is equivalent to constructing a rational function around each group of data points.

The program is a generalized procedure for producing an intuitively reasonable surface which takes specified values at specified data points.

In interpreting a SYMAP, one should recognize that it is perfectly reliable -- up to the accuracy of the data -- only at data points. Where the spacing of data points is relatively wide, or where the data value fluctuates rapidly from one data point to its neighbors, the SYMAP is less reliable and is best interpreted as a general picture. Only if the user is confident that his spacing of data points is finer than the scale of the phenomenon he is mapping, may he confidently assume that a SYMAP -- like any contour map based only on those data -- is valid.

2. SENSITIVITY AND ACCURACY OF INTERPOLATION

We were interested in measuring the "accuracy" of the SYMAP computer mapping program under various situations and seeing whether the interpolation algorithm might be improved upon. As a first step in this investigation, a study was made of the sensitivity of the interpolation to the number and placement of data points (14) for a series of approximations to a hypothetical surface.

Version V of SYMAP was used to produce a "reference map" of a surface used in previous studies.(15) 635 data points were inserted along a regular grid and at various "critical points" (16). Then two series of maps of the same surface were prepared. Series 1 used grids with 4, 9, 16, 64 and 272 data points; Series 2 used varying numbers of "critical points." In addition, all of the maps contained a band of 58 background values (outside of the map area) to provide for correct gradients at the edges of the map.

For each of the test maps in both series, the difference between test value and reference value at each of the data points (approximately 3000 locations) was computed, and the mean ($\bar{x}$) and standard deviation (s) of these differences were determined. (17) In general, as the number of data points increased, the mean and standard deviation of the differences tended toward 0. The means and standard deviations of the differences for Series 1 (gridded points) were found to be lower than those for maps in Series 2 (critical points). This last result was contrary to our expectations and suggested the need for some of the modifications in the interpolation algorithm described in the following sections.

2.1 SENSITIVITY TO TOTAL NUMBER OF DATA POINTS

To investigate the sensitivity of the interpolation algorithm to the number of data points the standard deviation of the differences was plotted against the number of data points, on log-log paper. (See example 2-1 and Table 2-1) It is clear that this plot is approximately linear with a slope of -0.5. From the plot of example 2-1 we may write the relationship:

$$S = 93.5/\sqrt{n}$$

Where S is the standard deviation of the differences,
N is the number of data points, and
93.5 is a constant whose units are the same as those of the original surface.

Since the original surface, with 635 data points, had a standard deviation of 45.6, the constant (93.5) represents about two standard deviations.

The above relationship can also be written in the form

Standard deviation of differences	x	number of points	=	constant

Since, on a regular grid, the distance between data points is inversely proportional to the number of data points, we may re-write the above relationship as:

standard deviation of differences	=	constant	x	distance between data points

It is interesting to note that, although approximately linear, the plot is slightly concave down. Therefore, a

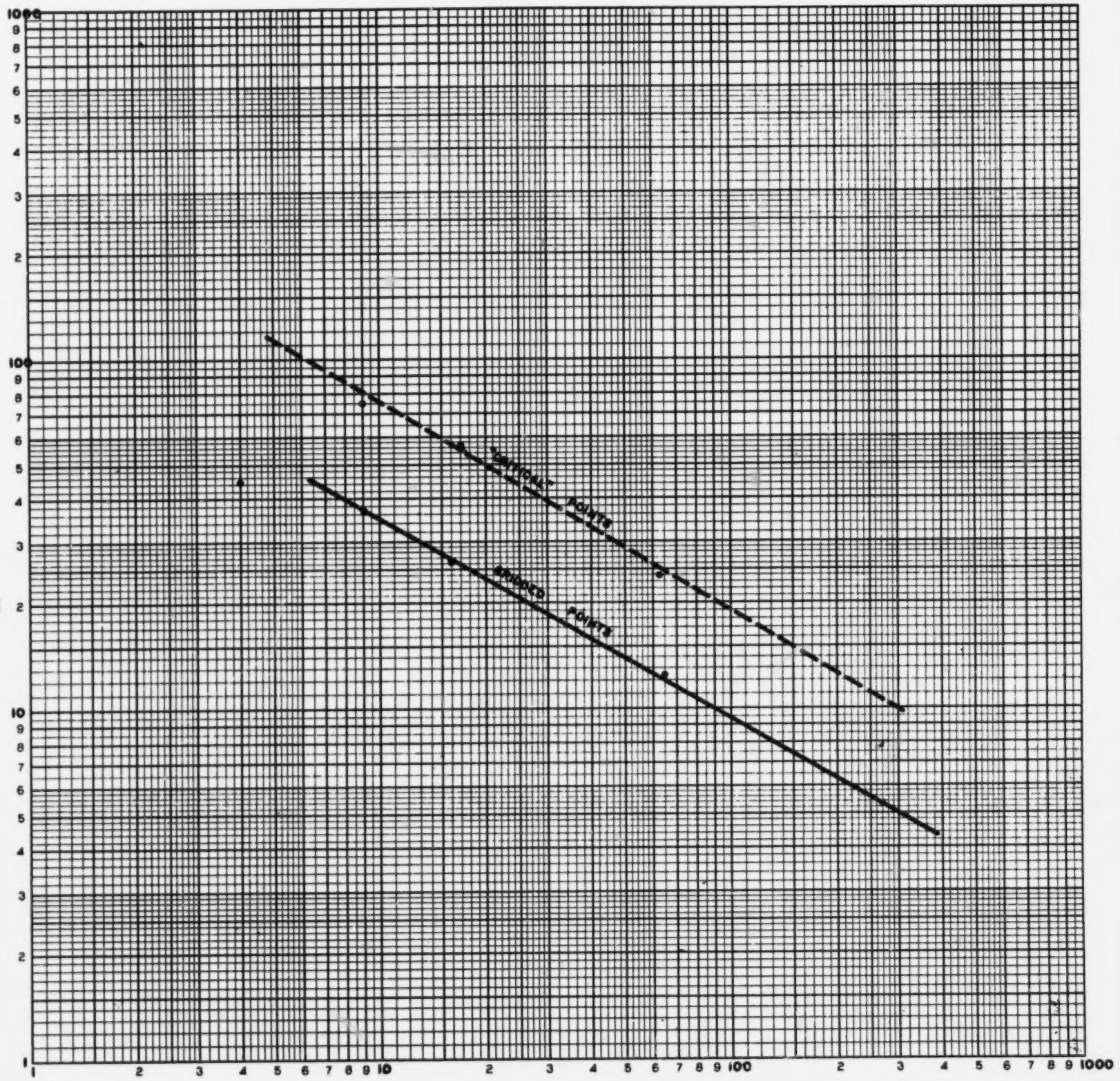
TABLE 2-1

SENSITIVITY AND ACCURACY OF INTERPOLATION:

Summary Statistics

	Original		Interpolation with 4 Points	
	$\bar{x}$	s	$\bar{x}$	s
Original Surface			106.81	45.62
Series 1: (Gridded Data Points 4 pts.	-33.14	45.15	-33.32	45.80
9	-15.05	37.63	-18.34	38.55
16	-1.89	26.39	-1.10	27.26
64	-0.22	12.72	0.17	13.39
272	-0.07	5.38	0.12	5.81
Series 2: (Selected Critical Points) 9 pts.	-30.22	76.15	-29.64	79.35
17	-9.61	58.57	-8.50	62.37
64	-1.27	29.06	-1.36	31.74

$\bar{x}$ Mean
s Standard deviation



Example 2-1

Old Interpolation

given increase in the number of data points results in a greater proportional reduction of the variance when the number of points is large, than when the number is small.

2.2. DISTRIBUTION OF ERROR

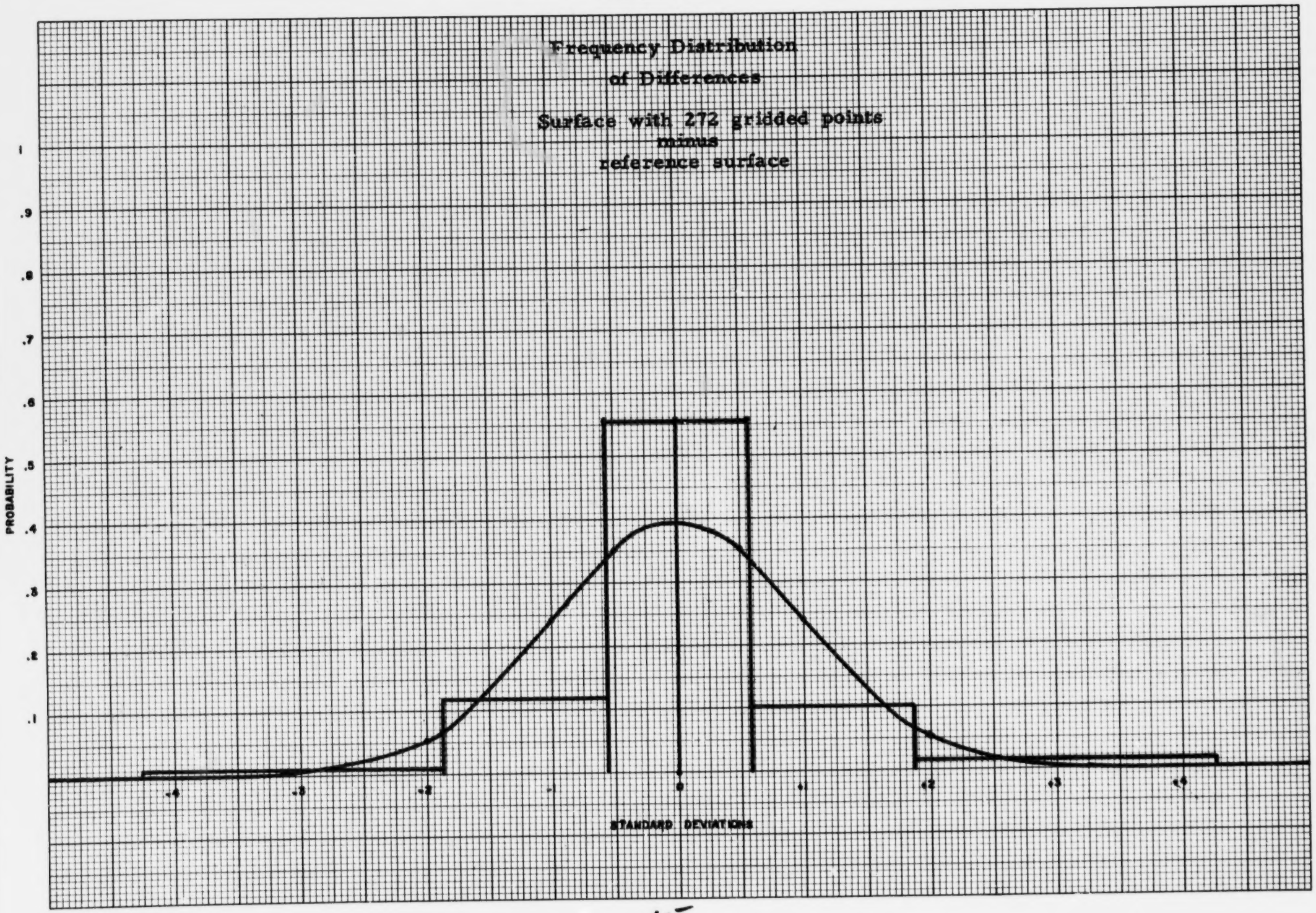
The test map with the smallest variation from the reference map was as expected, the map with the greatest number of data points (272 gridded points). Example 2-2 shows the frequency distribution of the differences, for this map, as a bar graph -- normalized by the standard deviation. A graph of the normal distribution with unit variance is superimposed. The relative frequency of very small differences (less than 0.6 standard deviations) is greater than the expected for the normal distribution. This is due in part to the fact that the surfaces match exactly, when 272 data points are used. The relative frequency of differences exceeding 2.5 standard deviations is also greater than expected under a normal distribution. From the foregoing, we can conclude that, in general, interpolation from a coarser grid of data points is a hit-or-miss process in which very good accuracy or a moderate chance of large inaccuracy are both possible.

2.3. IMPROVEMENTS IN THE INTERPOLATION ALGORITHM

2.3.1 Fixed Number of Data Points

Normally the algorithm chooses the number of data points to be used for interpolation at each character location of the map. For this test, however, the number of data points to be used for interpolation was fixed at 4 for every character location of the map. (18)

Example 2-2


Frequency Distribution
of Differences
Surface with 272 gridded points
minus
reference surface
PROBABILITY
1
.9
.8
.7
.6
.5
.4
.3
.2
.1
-4
-3
-2
-1
0
+1
+2
+3
+4
STANDARD DEVIATIONS

Although the surface was generally quite similar to the standard surface, it was somewhat smoother with fewer ripples. From the summary statistics of Table 2-1 and Example 2-3, it is seen that s is slightly (19) higher for this modification in interpolation. This factor must be weighed against the advantage of a "smoother" surface.

2.3.2. Inverse Distance Function

The standard interpolation algorithm is basically a weighting by inverse distance squared, with various correction terms. The algorithm was modified slightly to a weighting by inverse distance to the first power, with correction terms. Specifically, the equation

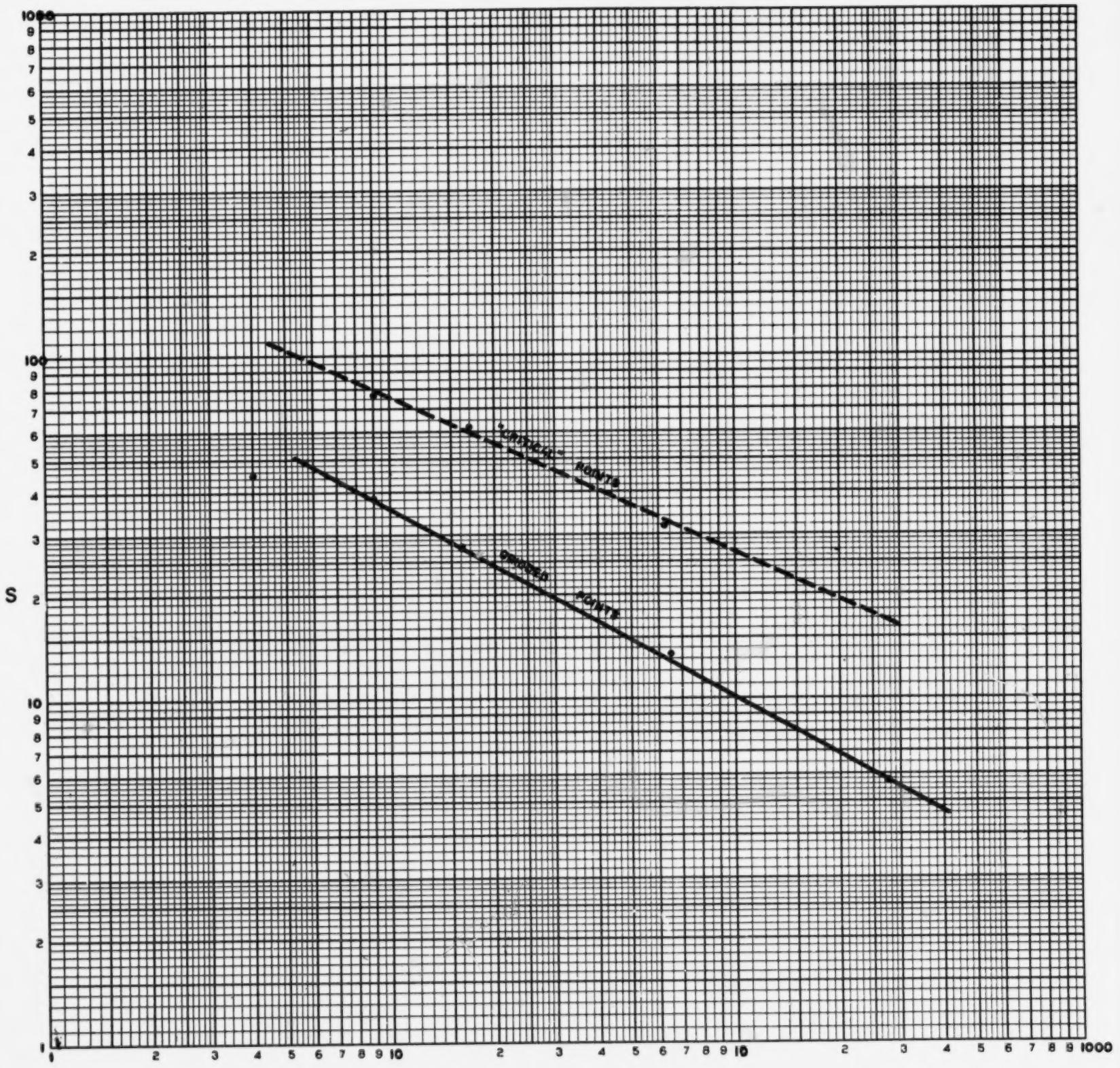
$$w_i = s_i^2 \ (1 + t_i) \quad \text{was changed to}$$

$$w_i = s_i \ (1 + t_i).$$

This new surface has the disadvantage of not having a continuous derivative at all data points. Since the function is evaluated at discrete character locations only, this was not found to be an hinderance. On the other hand, the slope at the data points is less constrained under the modified algorithm. The interpolated surface -- while still passing through all the data points -- is less complex, "smoother," and probably more reasonable for display purposes.

2.3.3. Comparison with Standard Procedure

We had planned to compare the interpolation procedure using both the above modifications with the standard procedure and evaluate the standard deviations of the differences. An error in the preparation of



Example 2-3

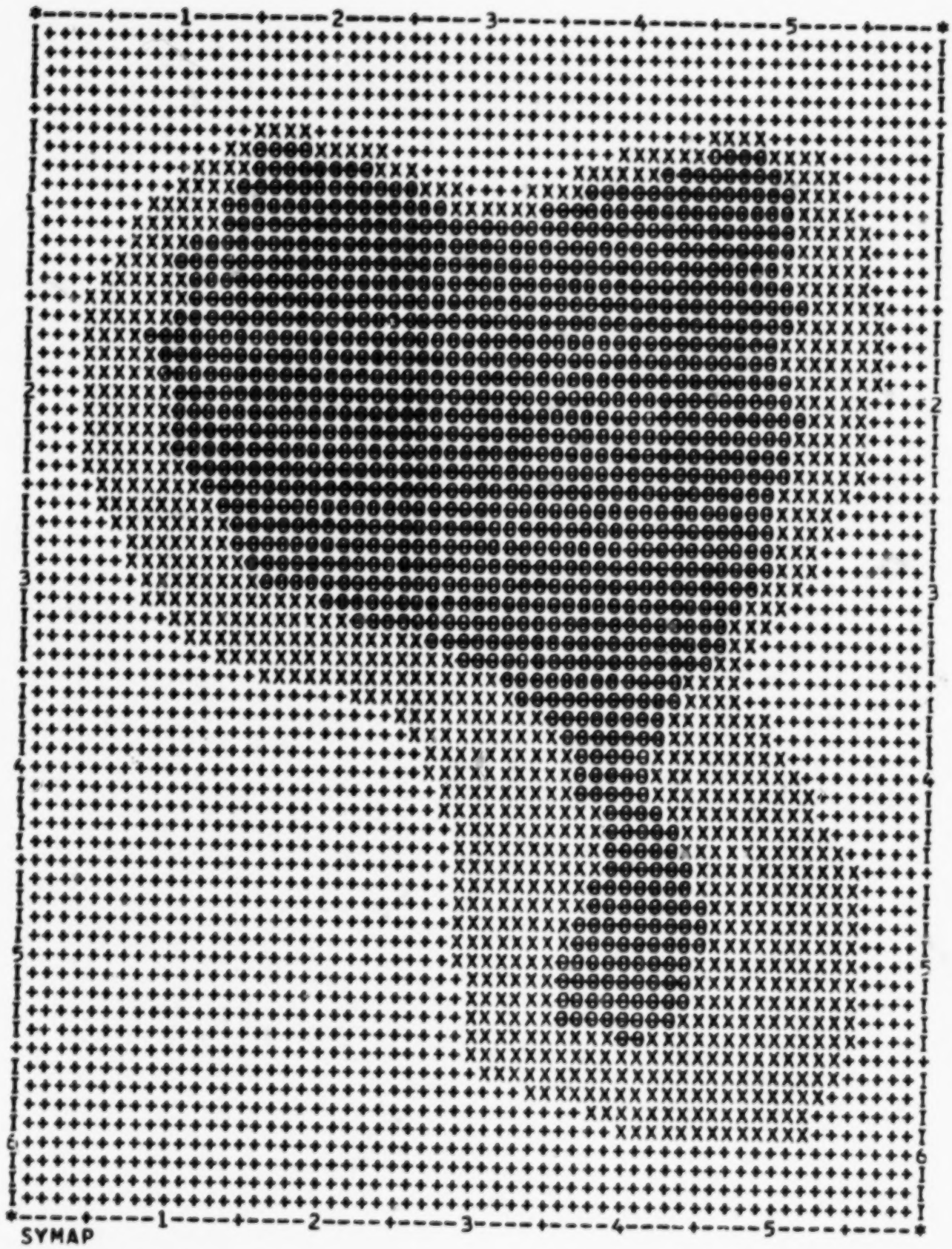
Interpolation with 4 Points

the reference surface made this quantitative comparison impossible. Instead we prepared Symaps (Examples 2-4 to 2-8) and Symvus (Examples 2-9 and 2-10) of the surfaces with grids of 4, 9, 16, and 272 data points. If we consider the last surface (272 points) as the "reference surface," it can be seen that as the number of points is increased, this reference surface is more closely approximated.

2.3.4. Conclusions

Investigations of the type described herein are important for the analysis of interpolated maps and for the planning of data collection schemes. At least for surfaces and the interpolation algorithm of the type considered here, data collected over a grid yields the greatest accuracy -- for a given number of points. And the "error" of an interpolated surface -- or the standard deviation of its differences from a reference surface -- is proportional to the distance between data points on the test surface.

Example 2-4



0.2 MINUTES FOR MAP

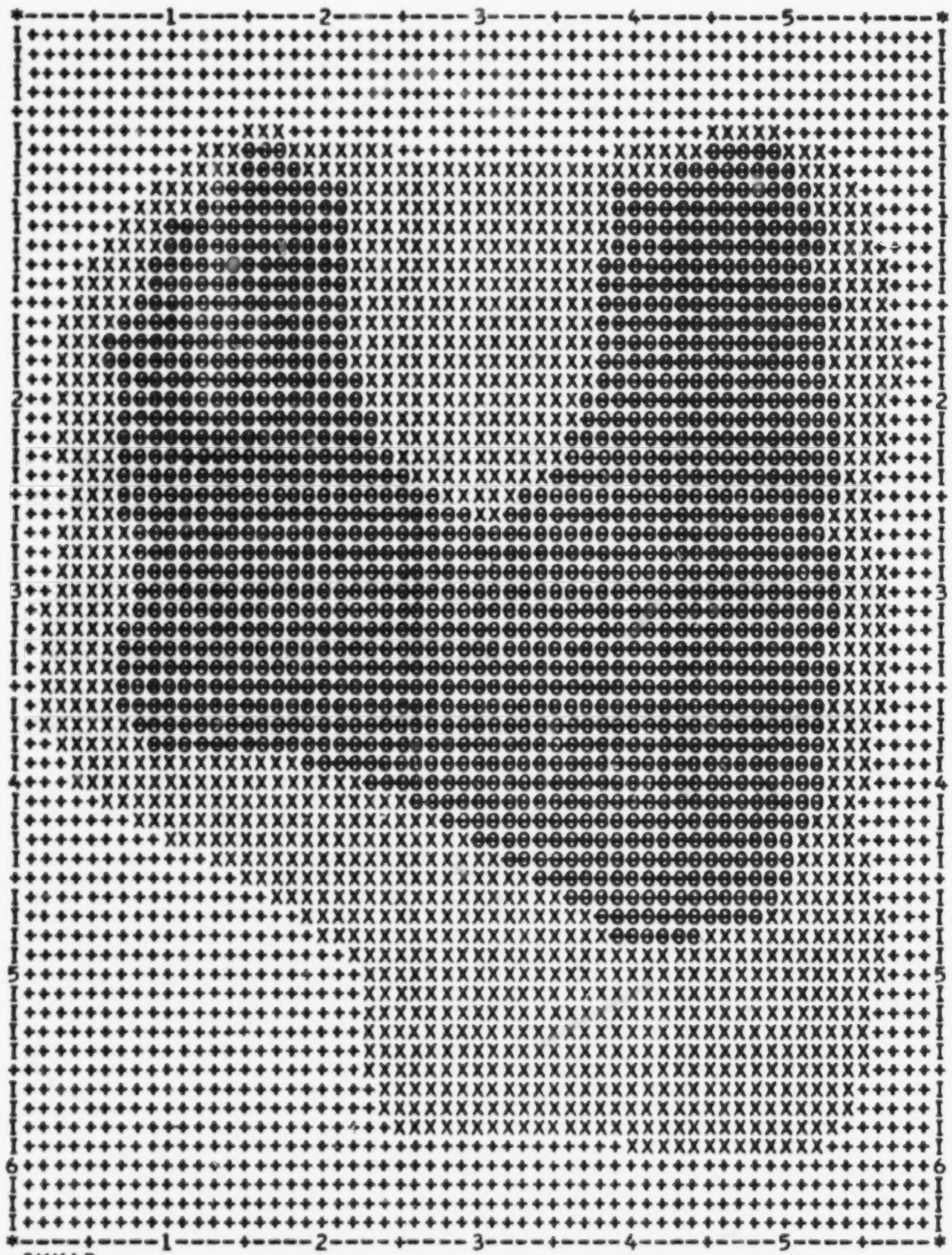
TIME = 672.7

THEORETICAL SURFACE

GRID ONE

(4)

Example 2-5



SYMAP

0.2 MINUTES FOR MAP

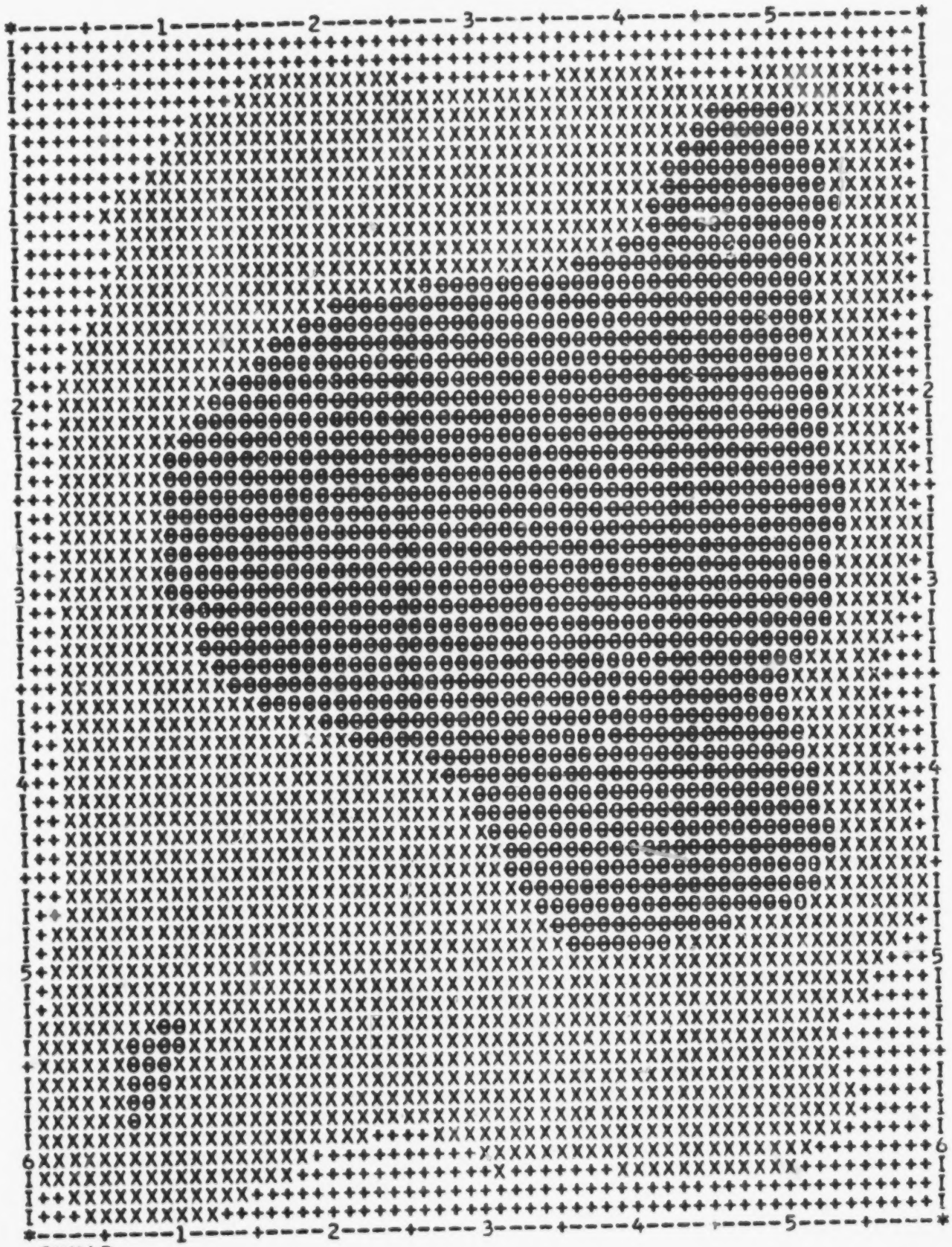
TIME = 672.9

THEORETICAL SURFACE

GRID ONE A

(9)

Example 2-6



SYMAP

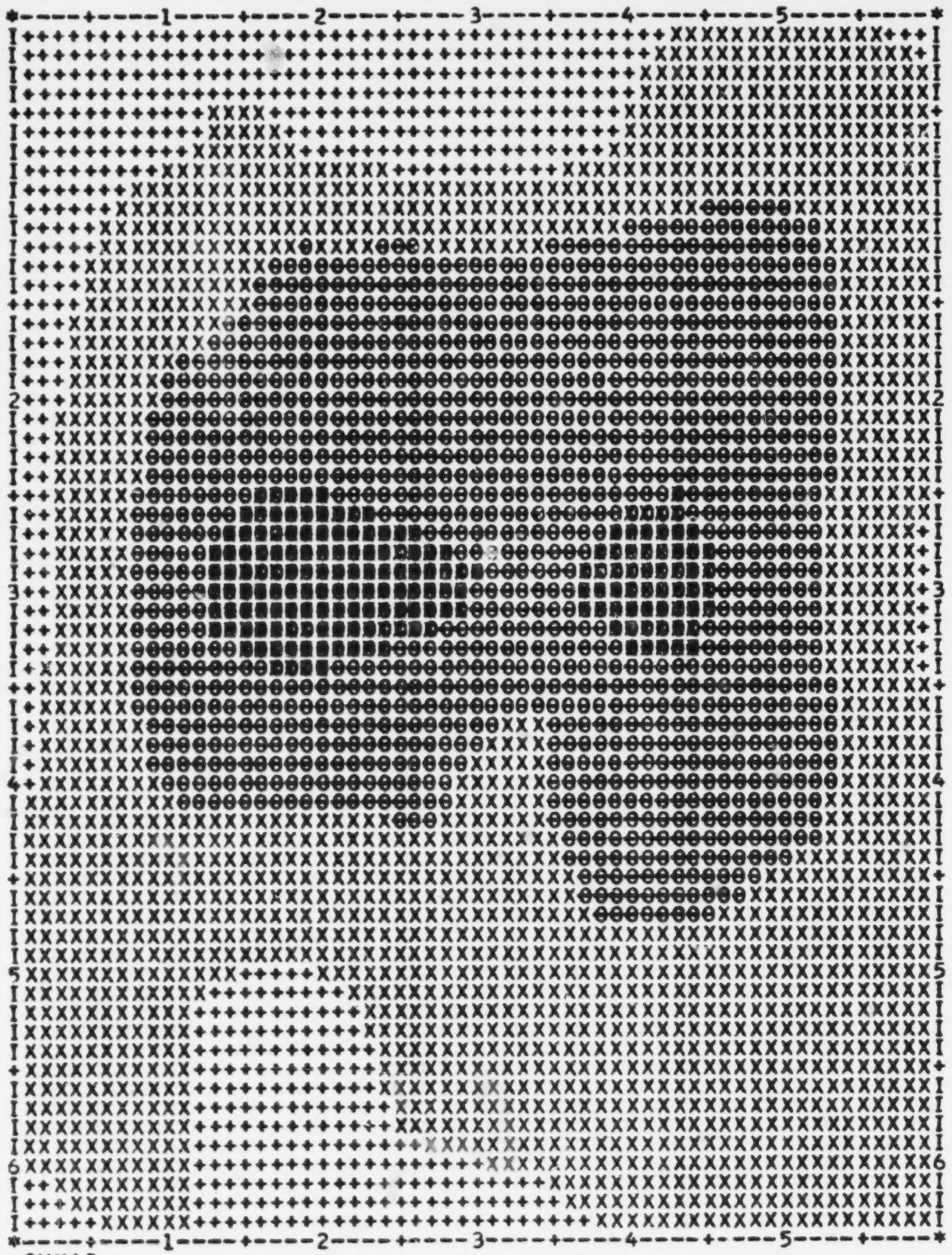
0.2 MINUTES FOR MAP

TIME = 673.1

THEORETICAL SURFACE

GRID TWO

Example 2-7



SYMAP

0.2 MINUTES FOR MAP

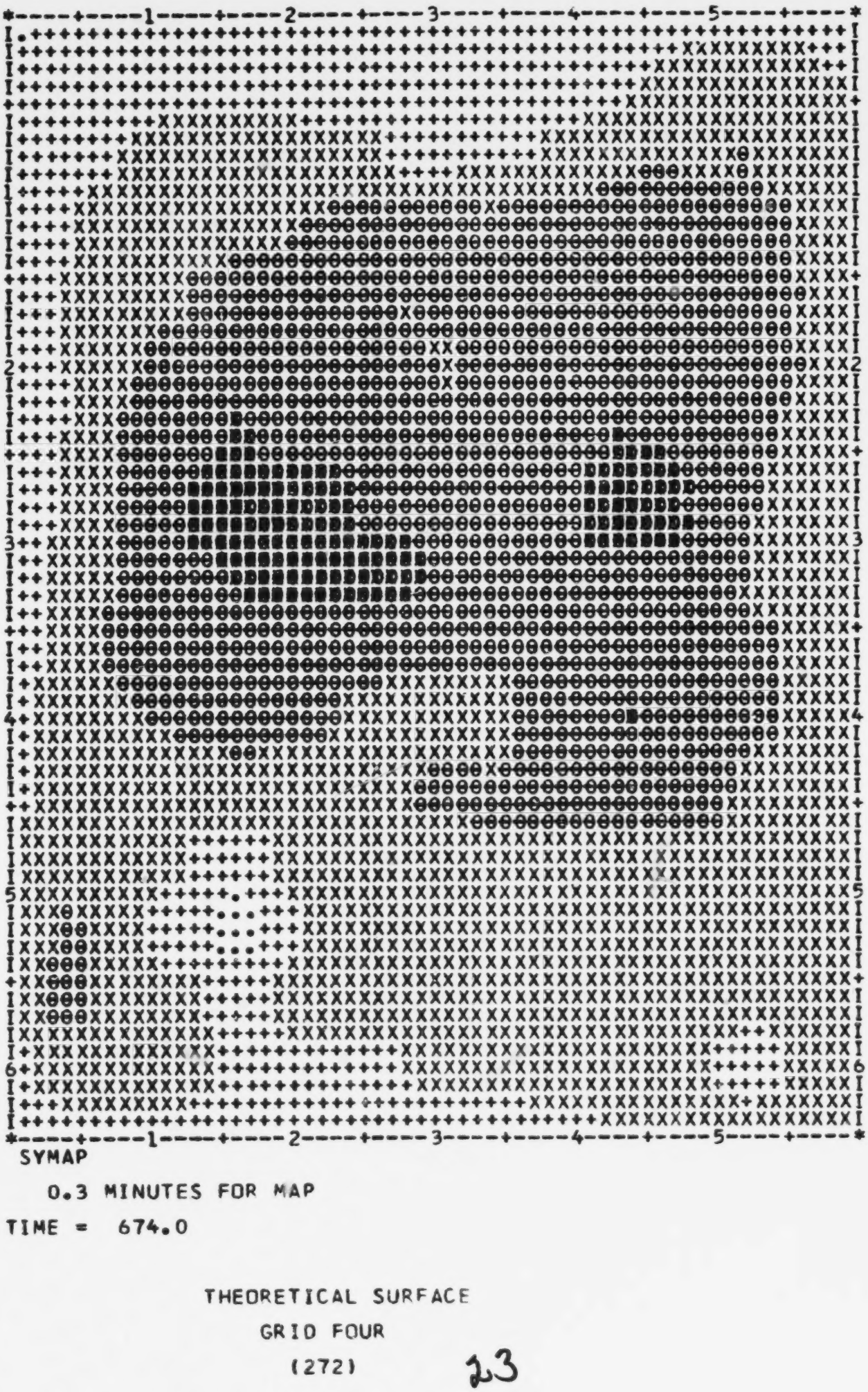
TIME = 673.5

THEORETICAL SURFACE

GRID THREE

(64)

Example 2-8



SYMAP

0.3 MINUTES FOR MAP

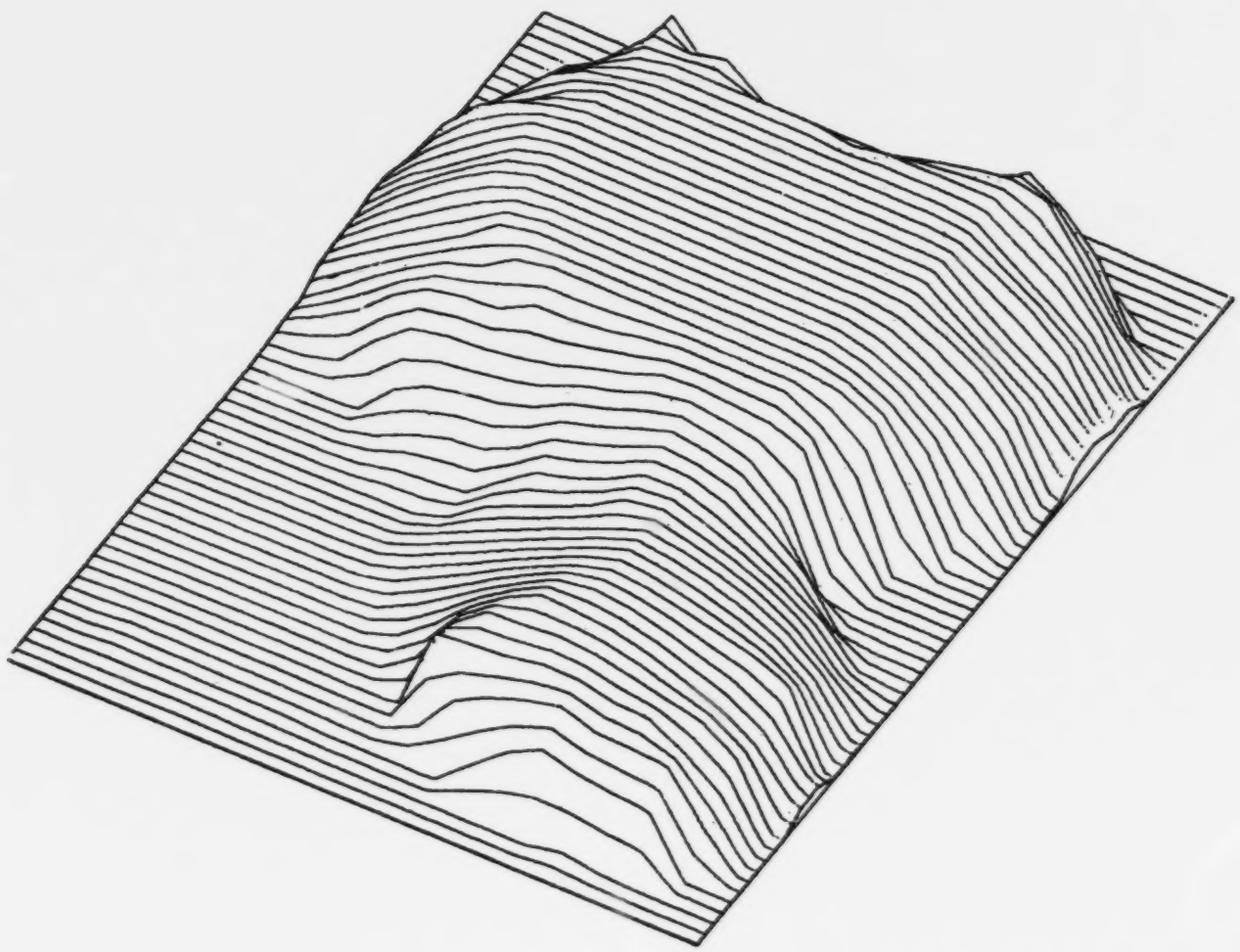
TIME = 674.0

THEORETICAL SURFACE

GRID FOUR

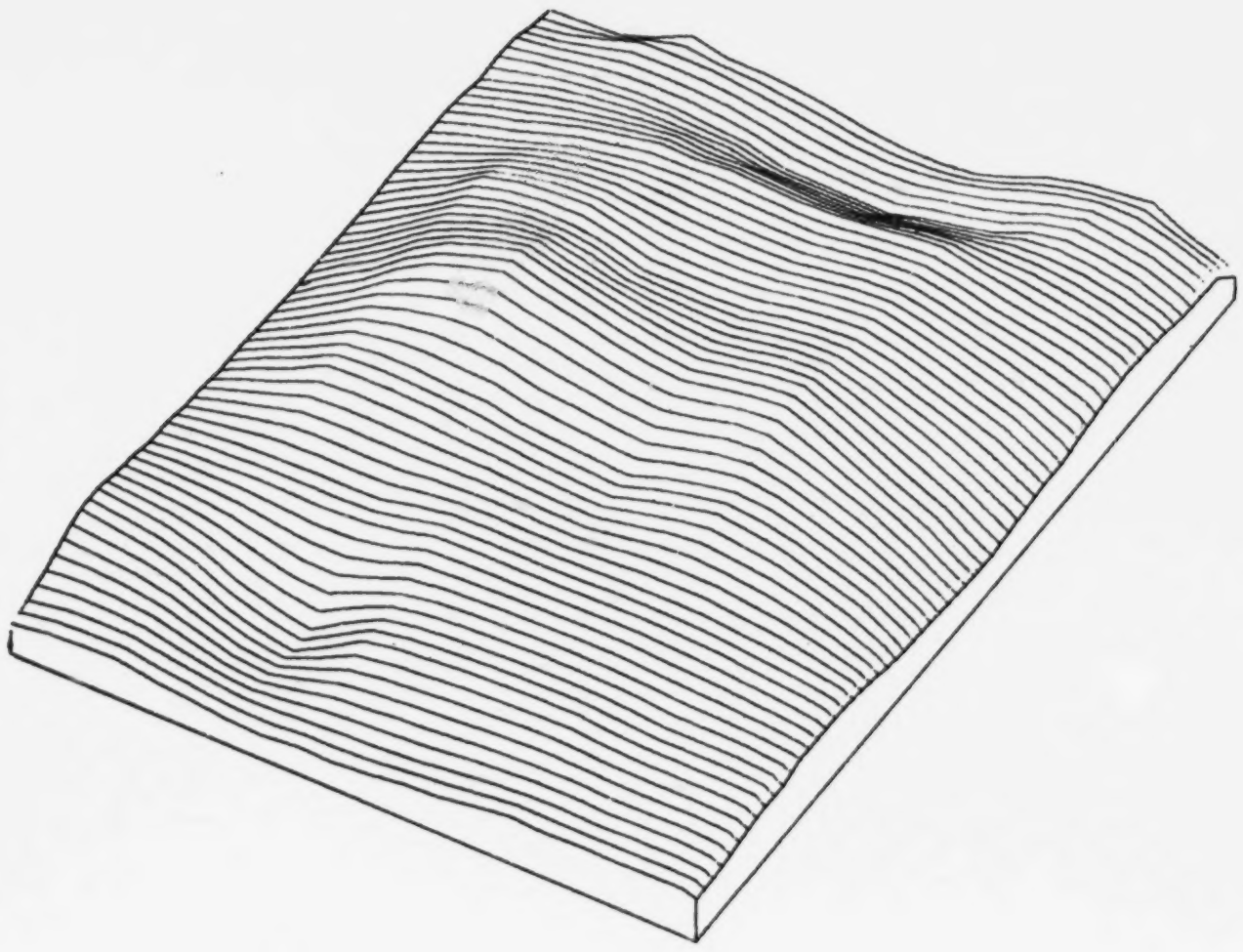
(272)

Example 2-9



```
MAP 2--4 DATA POINTS  (NEW INTERPOLATION)
06    05    69
AZIMUTH = 330              ALTITUDE = 45
WIDTH = 4.00              HEIGHT = 1.00
```

Example 2-10



```
MAP 5--272 DATA POINTS  (NEW INTERPOLATION)
06    05    69
AZIMUTH = 330              ALTITUDE = 45
WIDTH = 4.00              HEIGHT = 1.00
```

3. USE OF ELECTIVES TO AFFECT INTERPOLATION

The SYMAP interpolation algorithm is one of the more complex elements of the program. Basically, the program interpolates by taking a weighted average of the values of nearby data points. Although the form of this weighted average is fixed by the interpolation algorithm, certain parameters may be varied by the user. In this section the effect of setting these parameters to values other than their standard ones will be discussed.

3.1. PROPERTIES OF SYMAP INTERPOLATION

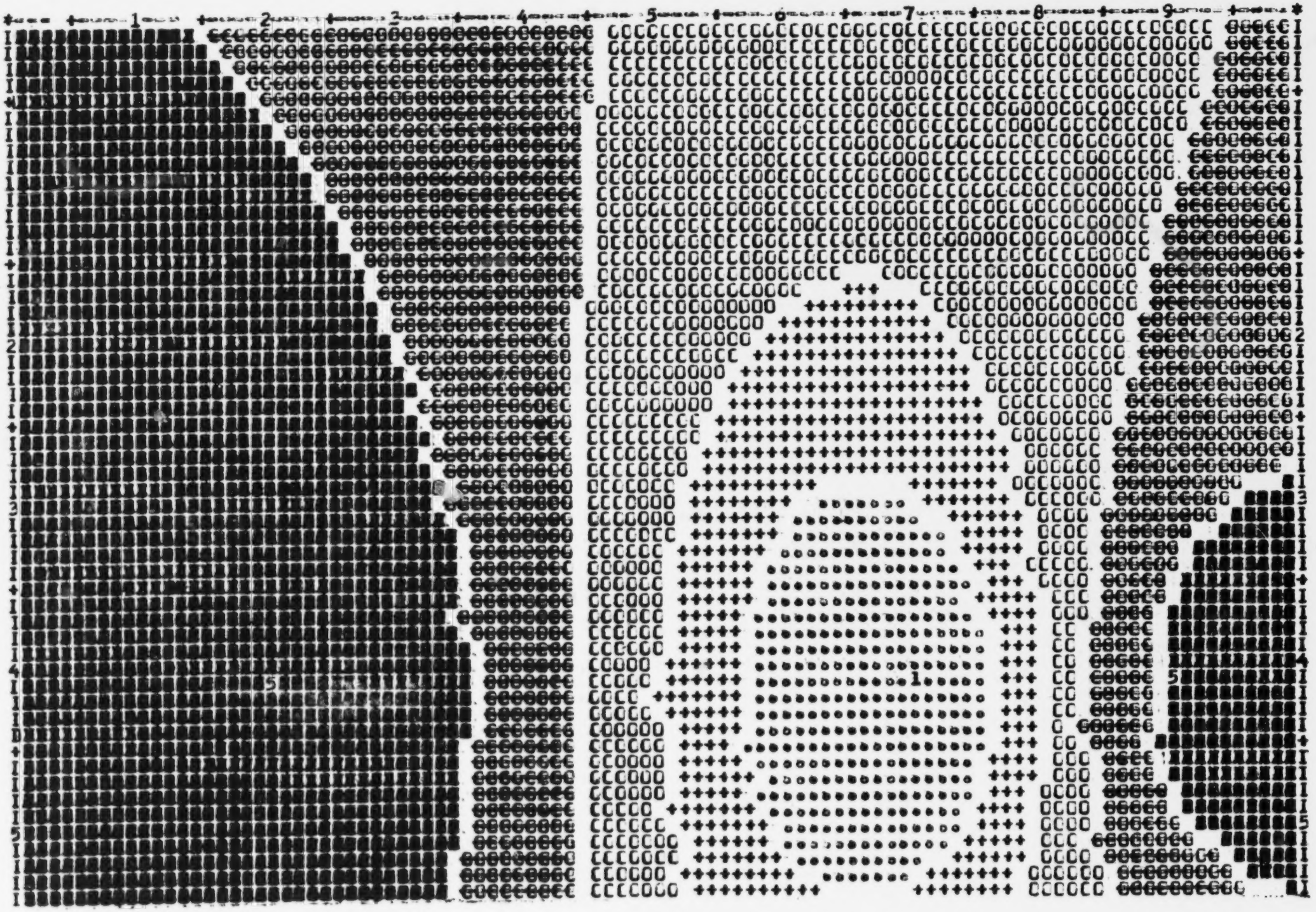
One of the significant features of the SYMAP algorithm is that it permits interpolation from points spaced unequally over a two-dimensional area. To understand how this interpolation works the simple one-dimensional configuration of Table 3-1 was used. (21)

TABLE 3-1

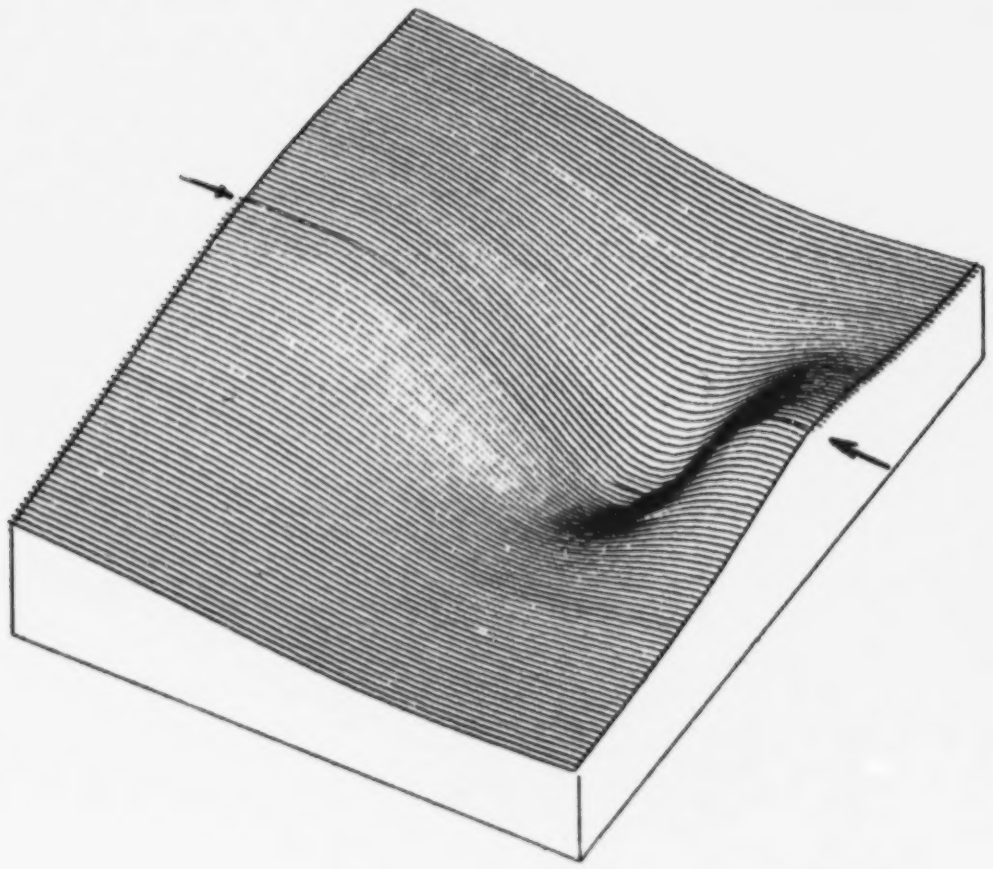
Co-Linear Data Points

Data Point Number	Coordinates y (down)	x (Across)	Data Point Value
1	4.1	2.0	9.0
2	4.1	7.0	2.0
3	4.1	9.0	8.0

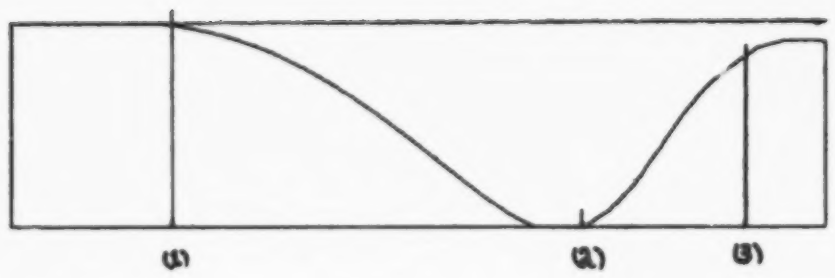
Example 3-1 shows a portion of a SYMAP based on these three points. It is clear from the map that the interpolation is not linear; the widths of the contour bands -- along the line connecting the three points -- are unequal. (22) Example 3-2 is a Symvu of the same surface. From this pair of maps, the behavior near the



Example 3-1
Interpolation Algorithm
3 Co-linear Data Points



Example 3-2
Interpolation Aglorithm
3 Co-linear Data Points
oblique view of surface



Example 3-3
Interpolation Algorithm
3 Co-linear Data Points
Profile

edges of the map becomes evident. At a control location that is approximately equidistant from all data points, the weighting of each of the data points is nearly equal because of the system of weighted averages that is employed. (23) And, at a control location that is far away from all the data points, the interpolated surface tends towards the average of the data points.

Example 3-3 is a slice along the axis of the three data points. It is really a Symvu of a single row from 0^{o} elevation. Again, the non-linearity of the interpolation is quite clear; the line between each pair of data points is not a straight line, but an S-curve. The curve is smooth with no sharp increases, even at the data points.

The behavior at the extremes of the value range is also interesting. The interpolation algorithm normally causes the surface to continue -- at every data point -- whatever trend, or slope, is indicated by the surrounding data points. In Example 3-3 the surface rises from data point 2 to 3 and continues to rise to the right of point 3. Here, as is often the case, interpolated maxima may be above the adjacent data point. (24) However, to insure that the overall range of the values on an interpolated surface is normally the same as the range of the original data values, a constraint is added:

> no interpolated value may exceed the overall maximum of the range of data nor fall below the overall minimum. (25)

In Example 3-3, although the profile shows the surface as rising from the right to data point 1, the

surface is constrained not to exceed this maximum. To the left of the data point the surface is a "flat plateau" at the level of this maximum value. Similarly, the surface is flat in the vicinity of data point 2, since it is constrained from going below this data value minimum. Data point 3 is a relative maximum data value but not an absolute maximum; therefore, the surface continues to rise to the right of this data point and the local maximum is higher than the data value.

3.2. VALUE EXTREMES

The property of interpolating to more extreme values than the original data values is under the user's control through elective 31, and a special use of electives 4 and 5. (26)

3.2.1. Plateau Effect

If the interpolated surface is to be used for Symvu display or certain other analyses, (27) then the flat areas around the extreme value data points are probably undesirable. To remove the plateau effect -- or areas of near zero slope -- electives 4 and 5 are employed as follows.

Although the data range extremes are normally the maximum and minimum data values, they may be set by the user. The flat area at the minimum data point can be removed by lowering the "value range minimum" (elective 4) by at least 10% of the data range. The flat area at the high data point can be removed by raising the "value range maximum" (elective 5) by at least 10%

of the data range. Through electives 6 and 7, levels and symbolism within this expanded range may then be set as desired.

3.2.2. Fractional Extrapolation

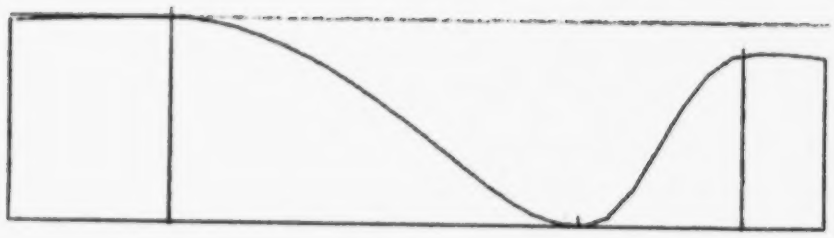
Elective 31 determines the fraction of the total data range that the algorithm will interpolate beyond a relative extreme data point. The normal value is 0.10, but users may wish to set this parameter to 0.0 so that extrema will always occur at data points. If this is done there will be no flat areas adjoining the extreme data points. However, a slope of zero occurs at every data point -- making it either an extremum or a point of inflection. Example 3-4 illustrates the effect of setting this elective to 0.0 in the previous example. (Example 3-3) As we move from data point 2 to either point 1 or 3, the surface first rises to a local maximum at the data point (as in Example 3-3), but then declines as we move past the data point.

In general the surface tends towards an average value in locations removed from the data points. For purposes of illustration, the fractional extrapolation was set to 0.2 in example 3-5. (28) If this elective is used, it should generally be to set the fractional extrapolation to 0.0.

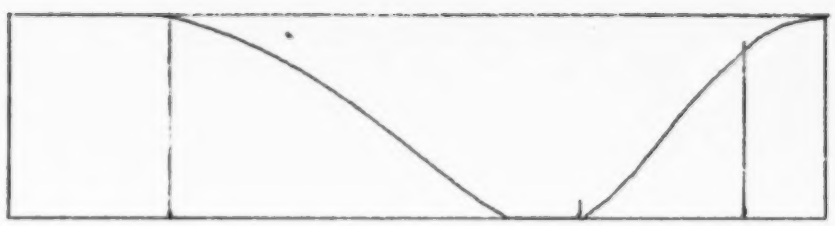
3.3. SEARCH RADIUS

3.3.1 Initial Radius

The initial search radius is an estimate of how far from any character location the computer must



Example 3-4
Interpolation Algorithm
3 Co-linear Data Points
Profile
Fractional Extrapolation 0.0



Example 3-5
Interpolation Algorithm
3 Co-linear Data Points
Profile
Fractional Extrapolation 0.2

search for "nearby" data points to use in interpolating the value at the character location. The overriding consideration, however, is that between 4 and 10 data points are required for interpolation. If this number is not obtained with the initial search radius on the first try, then the search radius is expanded (or contracted) until the requisite number of data points are included within the search radius. Since any initial radius may be modified by the program, this elective may have little effect on the final map.

This elective has several useful applications. If we are making a map in more than one section we would like to make the interpolation consistent from one set of data points to another. For example, suppose one wishes to make a map from 2000 data points. Since this number exceeds the maximum allowed in the SYMAP program (29) the map would be divided into two sections. Allowing for an overlap of some of the data points, each section would contain about 1100 data points. The search radius would normally be derived separately for each section, based on the density of data points. Taking the search radius of the first section, or map, as the standard, the search radius of the other map can be set to the first value on a later submission.

If we would like to fix the search radius at its maximum value initially, both electives 34 and 35 should be identical.

3.3.2. Maximum Radius

Elective 35 limits the distance over which the computer will search for nearby data points to use for

interpolation. The standard interpolation procedure will use all data points found within this maximum search radius, or the 10 closest data points if there are more than 10 points found.

This elective should be used whenever interpolation beyond some specified distance would not be meaningful. The average spacing between data points is often a good indicator of the maximum distance over which interpolation is valid. (30) Unless the data points nearest the map borders are known to be representative of a much wider area, it is reasonable to assume that extrapolation beyond the area enclosed by data points should be restricted to the average distance between data points. If there are no data points within the maximum search radius, then it is not possible to assign an interpolated value and the symbol for "no data" will be printed. (31)

This elective has proved useful for the mapping of air quality data for a metropolitan area. Often air monitoring networks have been set up in such a way that most of the monitoring stations are located in the more heavily polluted core area of the region. It would be misleading to interpolate from these central values to all the outlying areas. On the assumption that the original air quality monitoring network was sensibly established, the average data point spacing can be thought of as a reasonable estimate of the maximum distance over which are quality will not change significantly; this average spacing is, therefore, an appropriate value for electives 34 and 35. Figure C-1 (32) shows the application of this elective to maps of air quality for the St. Louis area. (33)

3.4 DETAIL OF SURFACE COMPUTATION

3.4.1. Number of Data Points

The number of data points on which interpolation is to be based is specified by elective 36. Normally, between 4 and 10 points are used, depending upon the search radius. Since a constantly changing number of data points is employed, the number of terms in the interpolation function also changes, and the resulting surface is full of ripples with unneccesarily jagged contour lines. (See Example 3-1) A somewhat smoother surface can be obtained by always using the same number of data points. If the minimum and maximum number of data points (34) are each set to 4.0, the interpolation should be smoother than with the standard setting of the elective. Users may wish to try this elective on maps with jagged contour lines.

3.4.2. Independent Computation

Elective 37 specifies that the interpolated value will be computed independently at every character location. This increases the computation time for the map by 100 to 200% over the standard method.(35) It may be useful in the following cases:

(i) proximal map;

(ii) with elective 35 whenever smooth edges are desired instead of a jagged outline; and

(iii) with barriers, when interpolated values adjacent to the barrier are printed and an especially sharp demarcation of the barrier is desired.

4. USE OF BARRIERS TO AFFECT INTERPOLATION

If we wish to prevent or inhibit interpolation from one area to another -- across some line dividing the areas -- we place a "barrier" to the interpolation along the dividing line. Barriers have been shown to be useful for numerous mapping situations, including:

i) planning -- in mapping demographic data, to indicate that one region is separate from another. For example, the population density in New York city may be so much higher than nearby areas that it bears little relation to the surrounding densities in the metropolitan area. Without the barrier, the very high densities would extend over a much larger area than is realistic;

ii) geology -- an oil company using SYMAP inserted barriers along known faults in maps of the depth of certain rock strata. With barriers, the geological maps closely resembled maps previously drawn by hand by cartographers trained in geology; and,

iii) air pollution -- to describe the air pollution surface for the Los Angeles area adequately, a barrier was used to show the effect of the San Gabriel mountains north of the city, where the air pollution levels decrease rather abruptly.

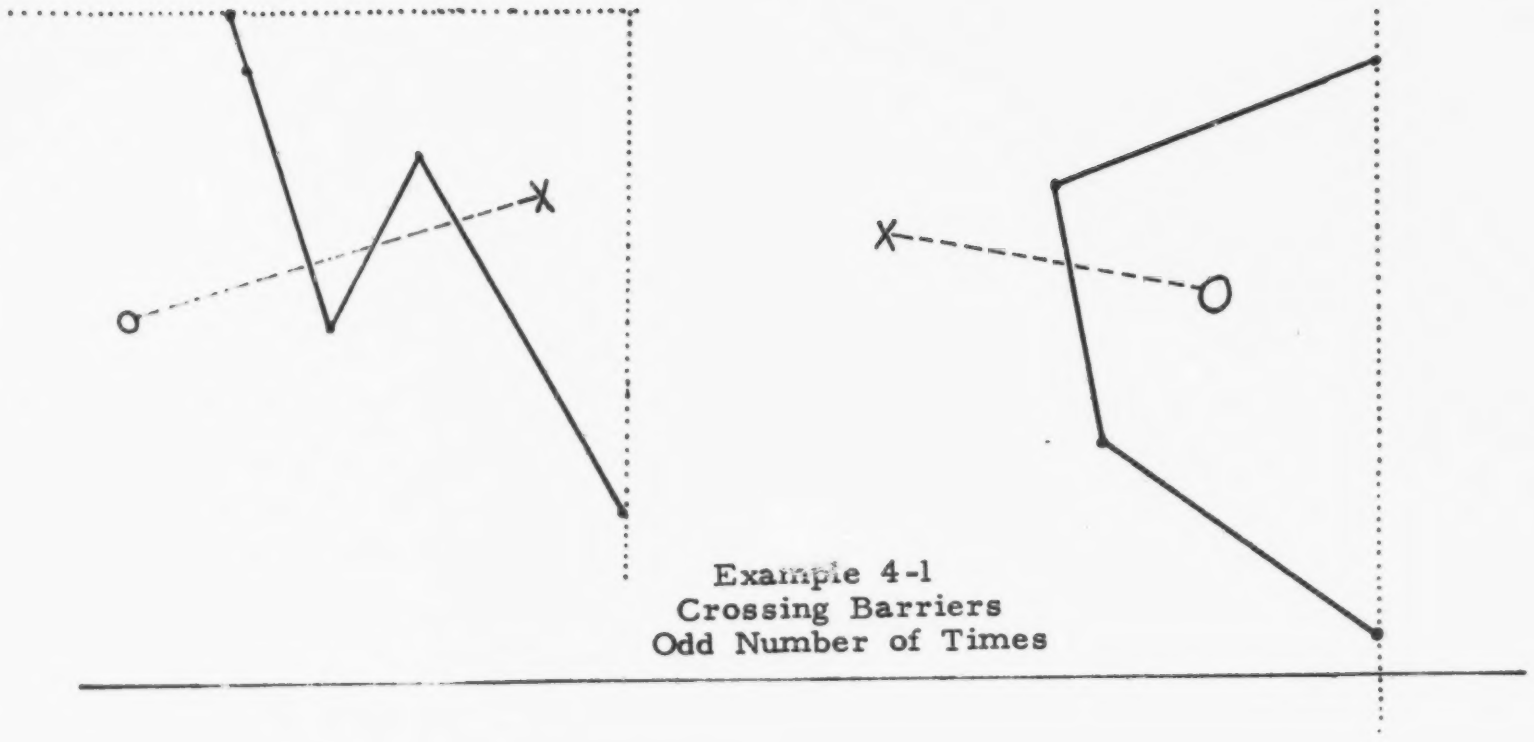
4.1 CROSSING BARRIERS

The barrier acts to make data points on the other side appear to be further away than they actually are. "Across" the barrier has a literal mathematical meaning: if the imaginary line from a control location to a data point crosses the barrier an odd number of times, then that point is across the barrier from the control location. If the line crosses the barrier an even number of times, then the barrier has no effect whatsoever on the interpolation from the data point to the character location under consideration. Example 4-1 shows examples in which a barrier is crossed an odd number of times, and therefore has an effect; Example 4-2 shows examples in which the barrier has no effect. (36)

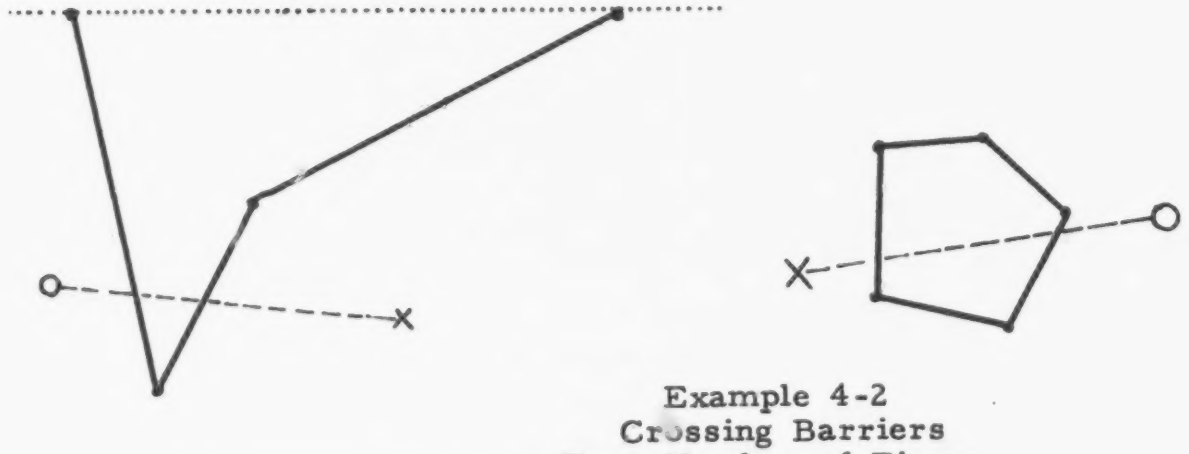
4.2. IMPERMEABLE BARRIERS

A barrier as used in SYMAP may be either "permeable" or "impermeable". Permeable barriers restrict interpolation across the barrier, but do not stop it. Impermeable barriers allow no interpolation across the barrier. (37)

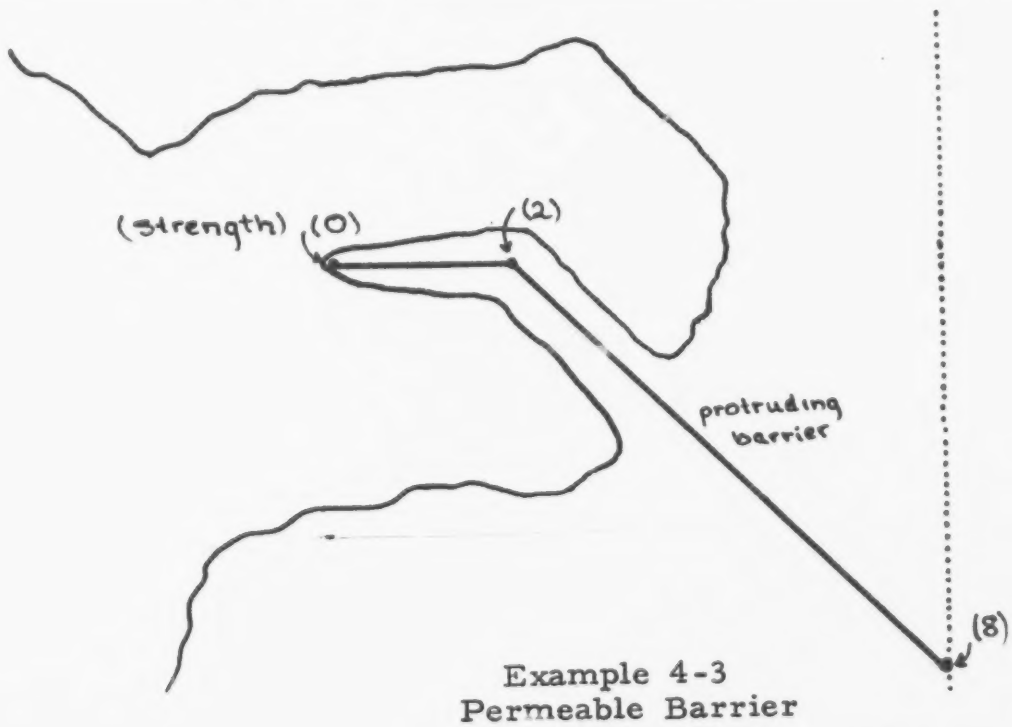
Of the two, impermeable barriers are the easier to understand. To be used correctly, an impermeable barrier must divide the study area into completely separate regions. Two types of impermeable barriers are thus possible: dividing line and closed loop.



Example 4-1
Crossing Barriers
Odd Number of Times



Example 4-2
Crossing Barriers
Even Number of Times



Example 4-3
Permeable Barrier
Protruding Type

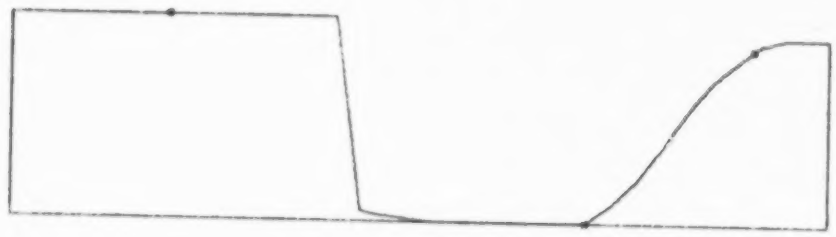
4.2.1. Dividing Line

A dividing line barrier begins at the map border and terminates at the map border, although it may change direction in between. (38) It completely divides the study area. For example, in a map of greater New York City, the Hudson River may extend up the center, separating New York and New Jersey so that interpolation, for many phenomena, should not occur from one state to the other. To convey this separation to the computer, a barrier would be placed from a point on the top border to a point on the bottom border of the map, passing down the center of the river.

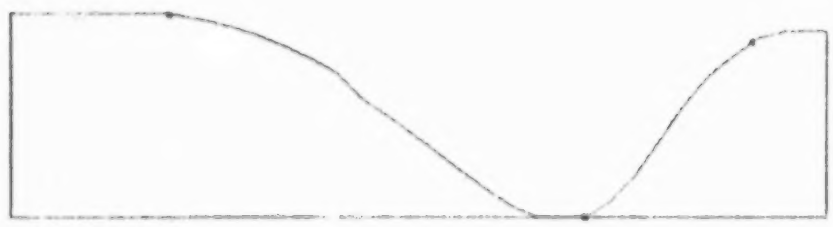
4.2.2. Closed Loop.

If one region is to be separated from another region which completely surrounds it, then a closed loop barrier should be placed around the interior region. A region such as an island -- or an isolated city -- which is not thought to have any relation to the surrounding area for the data under consideration, should be encircled by a closed loop barrier. For example, a closed loop barrier may be used to advantage in making a map of population density over the continental United States, based on data for forty-eight states and the District of Columbia. Because the District is a small, very densely populated region, data from it are not comparable with data from the states and should be considered separately.

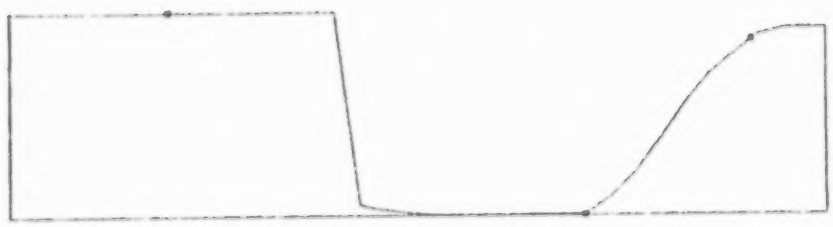
Example 4-4 shows a profile of a surface intersected by an impermeable dividing line barrier.(39)



Example 4-4
Impermeable Barrier
Dividing Line Type
Strength -1



Example 4-5
Permeable Barrier
Dividing Line Type
Strength 1



Example 4-6
Permeable Barrier
Dividing Line Type
Strength 5

Since an impermeable barrier has been used, we can assume that the surfaces on each side are logically separate. The left-hand surface contains only one data point, and is, therefore, flat. The right-hand surface contains two points; this surface remains flat and low over a large area in the center, since the left most data point has no effect to the right of the barrier.

4.3. PERMEABLE BARRIERS

Permeable barriers may be used to represent the inhibiting influence of bays, harbors, lakes, or other obstacles. Points across a lake, or at opposite ends of a bay, may still be thought to bear a relationship to each other, but the relationship would depend more on the overland distance. As in the case of impermeable barriers, several types are possible.

4.3.1. Dividing Line

The general rules are the same for permeable as for impermeable barriers. In the case of permeable barriers, the units (for barrier strength) are the same units in which data point locations are expressed. (40) For our example, we find -- from the original data point listing -- that five source units separate the data points which fall on either side of our barrier. Example 4-5 shows the effect of a permeable dividing line barrier of strength 1 -- there is no more than a slight change in the gradient at the barrier. (41) In Example 4-6 a barrier of

strength 5 was used and the profile is indistinguishable from the impermeable barrier case. Inspection of the numerical interpolated values, however, shows that data points at one side of the barrier do have an influence on the other side for a distance of a few character locations. If a permeable dividing line barrier is used, then the strength should generally be no greater than the average spacing between data points (42).

4.3.2. Protruding Barrier

The great advantage of a permeable barrier is that we may vary the strength along the barrier. For protruding barriers, this is an especially advantageous property.

A bay, harbor, or river mouth may be thought of as protruding into the study area, as Boston Harbor "protrudes" between Winthrop and South Boston, or as Cape Cod Bay "protrudes" between Provincetown and Plymouth. An impermeable barrier is not appropriate, because there are areas which provide a link between the tip of Cape Cod and the rest of Massachusetts, or between South Boston and Winthrop. The best solution in such cases is to use a protruding permeable barrier. The barrier should begin at the shoreline, continue into the bay or harbor, and end at the map border. Each vertex of this barrier will have a different strength, starting with a strength of 0 at the shoreline. The strength at the second vertex should be twice the length of the segment from the first to the second vertex, and so on for all the

vertices. (43) The strength at the final vertex on the map border will be roughly twice the total length of the barrier, in source units. The strength of zero at one end of the barrier means it has no effect at that point. The other strengths approximate the effect of having to double around the barrier to avoid crossing it. In other words the strength at each vertex of a protruding barrier is twice the distance from the first vertex. (See Example 4-3)

4.3.3. Other Configurations

Other configurations of permeable barriers are possible. In general, the strength at any terminal vertex within the study area should be zero. A long lake might need a barrier extending along its major axis, with zeroes at both ends and strengths increasing towards the center. An additional vertex would be needed at the center, having a strength equal to the length of the lake.

REFERENCES

(1) General references include:

Bengtsson, Bengt-Erik and Nordbeck, Stig, "Production of maps with the help of digital computers," Lund, Geografiska Institutionen, Vol. 40, 1964.

Birkhoff, Garrett, and DeBoor, Carl R., "Piecewise Polynomial Interpolation and Approximation", Approximations of Functions, Amsterdam, 1965.

Fisher, Howard T., "SYMAP, Version 3 FORTRAN Computer Program", Laboratory for Computer Graphics, Harvard University.

Krumbein, W. C., "Trend Surface Analysis of Contour-Type Maps with Irregular Contour Point Spacing", Journal of Geophysical Research, Vol. 64, No. 7, July 1959.

Tobler, Waldo R., "An Interpolation Algorithm for Geographical Data", Department of Geography, University of Michigan, (undated).

Several "electives" in the program and the provision for barriers permit the user to specify additional information. This is discussed in Section 4.

(2) Test maps with various exponents showed that the power 2 gave results most resembling intuitive hand contouring. Furthermore, it led to extremely simple computation.

(3) If Z_p is a function of the location p then all directional derivatives of Z_p vanish where p coincides with points 1, 2, 3, etc. This is shown by differentiating Z_p with respect to its coordinates.

(4) The user may instead specify the initial search radius with elective 34.

(5) Unless specified otherwise by the user (elective 36) the minimum allowed number of data points is 4, the maximum is 10. Then the average number of data points for interpolation is 7.

(6) So that no data points are dropped, the final search radius, R! should be greater than the distance to any of the data points selected. It is set to the distance to the next data point.

(7) In the SYMAP program, the cosines are easily computed as the scalar product of the vectors $\vec{iP}$ and $\vec{Pn}$, divided by the product of their magnitudes. The sum extends over all data points within the final search radius. When a data point's distance from P exceeds R'/3, the quantities $1/\overline{P1}$, $1/\overline{P2}$, etc., are replaced by a quadratic function of the distances $\overline{Pi}$. For $\overline{Pi} > r'/3$,

$$f(\overline{Pi}) = \frac{27}{4(R')3} \times (\overline{Pi}-R')^2.$$

This function is tangent to the initial weighting function, $1/\overline{Pj}$, at R'/3.

(8) This "fractional extrapolation" is normally 0.10. If the fractional extrapolation is set at 0. by elective 31, then extrema can occur only at data points and the surface is then level at every data point.

(9) Slightly above the average if these points indicated an upward trend in the direction considered, slightly below if they indicated a downward trend.

(10) Elective 37 overrides the use of grid points. A value is then computed at every location by the procedure described.

(11) This method leads to some smoothing at barriers or at the edge of flat-tone like areas produced by electives 35 or 36. In these cases option 37 is recommended. Normally, the effect of interpolating from grid points is trivial.

(12) Data points closer than 1 1/2 characters are ignored in determining slopes. With standard symbolism, the symbol "S" (superimposed) is printed at a character location in which more than 1 data point falls.

(13) If no barriers are included, effective distances and actual distances are identical.

(14) We were able to build upon work begun by Assistant Professor Thomas Peucker of the Department of Geography, Simon Fraser University, (Burnaby, B.C., Canada).

(15) The reference map was about six inches square; the surface had been used in previous studies by William Warntz.

(16) Pits, peaks, passes, and pales are the critical points or features of the surface. See the discussion by William Warntz in Volume Two, Report G on the Southern New England Study.

(17) Appendix 2 describes the computer program written to calculate these statistics.

(18) See the discussion of elective 36 in section 3 and in Volume Two, Report K.

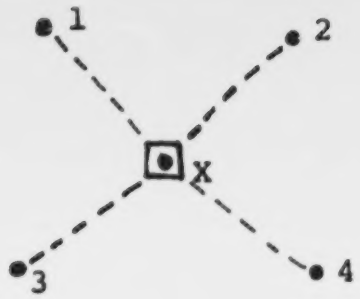
(19) The standard deviation is probably not significantly higher.

(20) See p. 520, col. 1 of Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly-Spaced Data," Proceedings of the 1968 Association for Computing Machinery National Conference, Brandon/Systems Press, Princeton, N. J. 1968.

(21) The user who desires to make any maps such as these with co-linear data points should be aware of the following: the interpolation algorithm will assign an initial search radius of 0 (since this estimate is based on the area occupied by the data points) and the program will stop. To overcome this, the user must compute the approximate spacing between data points ahead of time and specify this distance in elective 34.

(22) The value range of each level of symbolism is equal; therefore, if the interpolation were linear, the widths of the contour bands would be equal.

(23) For example, examine the two cases below. The interpolation to point X from data points 1, 2, 3, and 4 will result in the same value in both cases. (assuming the values of 1, 2, 3, 4 are the same). The interpolation takes into account only distances and values; directions, spacing, and clustering of points are not accounted for.



(24) The same holds true for interpolated minima.

(25) In Example 3-3 (produced by the Symvu program and output on a CALCOMP plotter) vertical lines were drawn to indicate the location of each of the data points and the points were numbered.

Point 1 (left side) is the data, or value range, maximum;
Point 2 (right center) is the minimum, white
Point 3 (extreme right) is a local maximum.

(26) This section, and those that follow, should be read in conjunction with the corresponding section of the Symap manual, Volume Two, Report K.

(27) For example, the "watersheds" program; see Volume Two, Report G, Appendix 2.

(28) This means that the extrapolated extrema can be up to 20% (of the data range) above or below the corresponding extreme data point.

(29) The version of Symap used here (Version V) normally has a limit somewhat greater than 1000 data points per map. Depending upon the core available, this number can be increased, at the rate of seven words per data point.

(30) This can be thought of as the scale of the smallest phenomena revealed by the set of data points with accuracy.

(31) "N" is standard, but it is often useful to change this symbol to blank by making column 27 of the Elective 7 symbolism cards blank; see Volume Two, Report K.

(32) See Volume Two, Report C, Appendix 3 Computer Maps; also examples may be found in Reports A, B, and D.

(33) On this map, the search radius was limited to 15,000 feet and the "no data" symbol changed to blank. The rather jagged outline of interpolated values is due to a computational shortcut. Basic interpolated values are derived only at every third character position horizontally, and every second one vertically. If any one of these basic character locations has no data, then the secondary locations derived from them will also show no data. Therefore, in Figure C-1, it can be seen that the indentations in the outline -- which should resemble arcs of circles -- are not in groups of one or two characters, but in multiples of three characters horizontally and two characters vertically. The jaggedness may be corrected through elective 37.

(34) Columns 11-20 and 21-30, respectively.

(35) The standard method -- see reference 33 above -- derives certain values from previously computed ones, and thus involves less computation time. Only one out of every six values are independently computed with the standard method.

(36) On these figures, the "x" denotes the character location, the "o" the data point.

(37) Whenever impermeable barriers are possible, permeable barriers may be substituted. In general, permeable barriers may be thought of as hurdles whose height above any point is the strength of the barrier at that point. Interpolation across the barrier would be affected by the added distance in jumping the hurdle. Impermeable barriers would be insurmountable hurdles.

(38) The dividing line barrier may terminate at the same or different border.

(39) Its "strength" is a special code value of -1.0 -- this is used with all impermeable barriers.

(40) However, if data point locations are in rows and columns, the units for barrier strength should be inches.

(41) Note that the strength of 1 is in the same source units as the data points.

(42) This distance -- the spacing between data points -- is approximately the "standard search radius".

(43) e.g. the strength at vertex 3 is twice the distance from 1 to 2 plus twice the distance from 2 to 3.

APPENDIX 1

SYMAP COMPARED TO OTHER METHODS OF INTERPOLATION

The problem of interpolation has been of long and widespread interest, and it is instructive to mention several representative methods in order to evaluate SYMAP.[1]

Interpolation algorithms have been programmed to handle a special case of two dimensional interpolation, in which the given points form a regular lattice. Among them are several which fit a polynomial to data points. A program in the computer SHARE library developed at M.I.T. uses the nearest four or sixteen points. A method by Tobler[2] uses the nearest six points. An algorithm by Birkhoff, de Boor, et al. takes both the values and the derivatives as given. H. Fisher at Harvard used a hyperbolic paraboloid through the surrounding four data points in the initial versions of SYMAP. An algorithm for computer mapping developed by Bengtsson and Nordbeck recommends partitioning the domain into triangles with data points as vertices, and passing a plane through the values at each of the data points.

When the data points do not form a regular grid, the problem becomes much more complex. None of the foregoing methods has been generalized satisfactorily, as each requires partitioning the domain into triangles or quadrilaterals. The difficulty comes in that the partition is not unique, even with such an additional constraint as minimizing the perimeter of all the polygons.[3]

A regular grid may be passed over the irregular data points and each data point assigned to the nearest grid call. This has been suggested by Bengtsson and Nordbeck. W. Tobler extended the method of piecewise polynomial functions, always fitting a double quadratic function to the six nearest points, and averaging its value with the nearest data point. Though the resulting function is not continuous at the junctions of polynomials, its simplicity and generality have made it useful in contour mapping.

Related approaches include the use of trend surfaces, for which programs have been written by Krumbein, or special purpose programs for which several types of data are available, as in meteorology or oil geology.

The SYMAP Version 4 algorithm, using piecewise rational functions, has been developed to provide a continuously differentiable function for both regular and irregular data point arrangements. The number of data points used in constructing each rational function is under user control.[4] It does not assign data points to a regular grid and it has the ability to extrapolate to points beyond the polygon enclosed by the data points. Unlike other methods, the SYMAP algorithm permits the inclusion of barriers.

REFERENCES

(1) The SYMAP program is an example of interpolating values for a function with two variables, from certain discrete points. Coordinate pairs (y_i, y_i) with

corresponding values Z_i are given for a finite set of discrete points and a function $f(x, y)$ is desired, satisfying, among other properties, $f(x_i, y_i) = Z_i$. Usually the interpolating function should be continuous or differentiable.

(2) See Reference 1 of the text for this and following citations.

(3) Further, as computer algorithms for partitioning the area had not been incorporated, this step had to be done manually.

(4) Nearest neighbor or "proximal" maps are included as a special case.

APPENDIX 2

MAP COMPARISON PROGRAM

The purpose of the program is to read a set of map profiles from a tape with the format used by SYMAP elective 21. Beginning with the second map, each map is compared with the first. If the first is assumed to be an "accurate" representation of a given surface, the program provides a measure of the accuracy of the subsequent maps.

In addition, a tape with profiles for maps of the differences between the locations on the first map and those on the other maps may be obtained.

The printed output from this program includes:

(i) mean, mean square, variance, and standard deviation of the reference surface;

(ii) number of rows and columns on each map;

(iii) mean, mean square, variance, and standard deviation of residuals (obtained by subtracting values at all locations on the first map from those on each of the other maps) and of their absolute values; and,

(iv) chi-square for each map, calculated using the values of the first map as "expected" values.

In addition there is a profile of the map of residuals, written on tape using the same format as SYMAP elective 21.

The program was written in FORTRAN IV for an IBM 7094. All maps must be of the same size and the number of rows and columns must be no more than 200. A standard FORTRAN deck setup is used. The tape with the map profiles must be mounted on FORTRAN symbolic unit 8, and the tape on which map profiles are to be written, on unit 9.

The following data cards are necessary:

(i) First card.

cols.	1-4	Number of maps on tape
	5-8	1 (integer) if tape desired

(ii) Cards with titles of second and subsequent maps in cols. 1-72. None are to be included for the first map. Three must be included for each of the other maps, although blank cards may be used.

Sample input and output for the map comparison program are shown on the next four pages. Here two surfaces from section 2 (grid--4 points, and grid--9 points) are compared to the reference surface. In the first run of this program, eight maps, with 63 rows and 59 columns each, were compared with a reference map. 2.9 minutes of computer time were required on an IBM 7094.

MAP COMPARISON PROGRAM

Sample Input:

```
C      MAP COMPARISON PROGRAM
C
   10 FORMAT(2I4)
   20 FORMAT(12A6)
   30 FORMAT(1H1,3HMAP,I3//)
   40 FORMAT(1X,12A6)
   50 FORMAT(16HONUMBER OF ROWS=,I3/19H NUMBER OF COLUMNS=,I3/20H DEGRE
     CS OF FREEDOM=,I6)
   60 FORMAT(6H MEAN=,F12.3/13H MEAN SQUARE=,F12.3/10H VARIANCE=,F12.3/
     C20H STANDARD DEVIATION=,F12.3/9H MAXIMUM=,F12.3/9H MINIMUM=,F12.3
   70 FORMAT(10HORESIDUALS)
   80 FORMAT(28HOABSOLUTE VALUE OF RESIDUALS)
   90 FORMAT(12HOCHI-SQUARE=,F12.3)
      DIMENSION AMAP(24),BMAP(200),    CMAP(26),DMAP(200),R(200),
     CTITLE(12)
      LOGICAL EOF
      ORMAX=-100000.0
      ORMIN=100000.0
      SUM=0.0
      SUMSQ=0.0
C
C      READ NUMBER OF MAPS
C
      READ(5,10) N,IMAP
      REWIND 1
      REWIND 8
      REWIND 9
C
C      READ NUMBER OF ROWS AND COLUMNS
C      IROW=NUMBER OF ROWS, ICOL=NUMBER OF COLUMNS
C
      READ(8) IROW,ICOL,(AMAP(J),J=1,20)
      IROW=IROW-1
      ICOL=ICOL-1
C
C      COMPUTE TOTAL NUMBER OF LOCATIONS, DEGREES OF FREEDOM
C
      KDF=IROW*ICOL
      PROD=KDF
      KDF=KDF-1
      DF=KDF
C
C      READ FIRST MAP
C
```

```
      DO 100 I=1,IROW
      READ(8) (BMAP(J),J=1,ICOL)
      DO 101 J=1,ICOL
      ORMIN=AMIN1(ORMIN,BMAP(J))
      ORMAX=AMAX1(ORMAX,BMAP(J))
      SUM=SUM+BMAP(J)/PROD
  101 SUMSQ=SUMSQ+BMAP(J)**2
  100 WRITE(1) (BMAP(J),J=1,ICOL)
      ASUMSQ=SUMSQ/PROD
      ORVAR=(SUMSQ-PROD*SUM**2)/DF
      ORSTDV=SQRT(ORVAR)
      WRITE(6,5)
    5 FORMAT(13H1ORIGINAL MAP//)
      WRITE(6,60) SUM,ASUMSQ,ORVAR,ORSTDV,ORMAX,ORMIN
      N=N-1
      DO 1000 K=1,N
      REWIND 1
C
C     READ NEXT MAP
C
      IF(EOF(8)) GO TO 150
  150 READ(8) (CMAP(L),L=1,22)
      IF(IMAP.EQ.1) WRITE(9) (CMAP(L),L=1,22)
      SR=0.0
      SR2=0.0
      ASR=0.0
      C=0.0
      RMAX=-10000000.0
      ARMAX=-10000000.0
      RMIN=10000000.0
      ARMIN=10000000.0
      DO 200 I=1,IROW
      READ(8) (DMAP(J),J=1,ICOL)
      READ(1) (BMAP(J),J=1,ICOL)
C
C     COMPUTE RESIDUALS AND WRITE ON TAPE
C
      DO 250 J=1,ICOL
      R(J)=DMAP(J)-BMAP(J)
      R1=ABS(R(J))
      RMAX=AMAX1(R(J),RMAX)
      ARMAX=AMAX1(R1,ARMAX)
      RMIN=AMIN1(R(J),RMIN)
      ARMIN=AMIN1(R1,ARMIN)
      R2=R1**2
      SR=SR+R(J)
      ASR=ASR+R1
      SR2=SR2+R2
  250 C=C+R2/BMAP(J)
      IF(IMAP.EQ.1) WRITE(9) (R(J),J=1,ICOL)
  200 CONTINUE
      IF(IMAP.NE.1) GO TO 201
      END FILE 9
  201 M=K+1
      WRITE(6,30) M
C
C     READ AND PRINT TITLE OF MAP
C
```

```
      DO 400 KM=1,3
      READ(5,20) (TITLE(J),J=1,12)
  400 WRITE(6,40) (TITLE(J),J=1,12)
C
C     COMPUTE STATISTICS
C     KDF=DEGREES OF FREEDOM, AMEAN=MEAN OF RESIDUALS, A2MEAN=MEAN
C     SQUARE OF RESIDUALS, VAR=VARIANCE OF RESIDUALS, STDEV=STANDARD
C     DEVIATION OF RESIDUALS, AVMEAN=MEAN ABSOLUTE VALUE OF RESIDUALS,
C     AVAR=VARIANCE OF ABSOLUTE VALUE OF RESIDUALS, ASTDEV=STANDARD
C     DEVIATION OF ABSOLUTE VALUE OF RESIDUALS, C=CHI-SQUARE
C
      AMEAN=SR/PROD
      A2MEAN=SR2/PROD
      VAR=(SR2-PROD*AMEAN**2)/DF
      STDEV=SQRT(VAR)
      AVMEAN=ASR/PROD
      AVAR=(SR2-PROD*AVMEAN**2)/DF
      ASTDEV=SQRT(AVAR)
      WRITE(6,50) IROW,ICOL,KDF
      WRITE(6,70)
      WRITE(6,60) AMEAN,A2MEAN,VAR,STDEV,RMAX,RMIN
      WRITE(6,80)
      WRITE(6,60) AVMEAN,A2MEAN,AVAR,ASTDEV,ARMAX,ARMIN
      WRITE(6,90) C
 1000 CONTINUE
      REWIND 9
      REWIND 8
      STOP
      END
```

Sample Output

```
ORIGINAL MAP

MEAN=      106.814
MEAN SQUARE=    13489.878
VARIANCE=     2081.148
STANDARD DEVIATION=        45.620
MAXIMUM=       249.633
MINIMUM=       -67.958
```

MAP 2

GRID--FOUR POINTS

NUMBER OF ROWS= 63
NUMBER OF COLUMNS= 59
DEGREES OF FREEDOM= 3716

RESIDUALS
MEAN= -33.317
MEAN SQUARE= 3206.920
VARIANCE= 2097.432
STANDARD DEVIATION= 45.798
MAXIMUM= 83.003
MINIMUM= -154.723

ABSOLUTE VALUE OF RESIDUALS
MEAN= 45.057
MEAN SQUARE= 3206.920
VARIANCE= 1177.060
STANDARD DEVIATION= 34.308
MAXIMUM= 154.723
MINIMUM= 0.000

CHI-SQUARE= 126490.775

MAP 3

GRID--NINE POINTS

NUMBER OF ROWS= 63
NUMBER OF COLUMNS= 59
DEGREES OF FREEDOM= 3716

RESIDUALS
MEAN= -18.338
MEAN SQUARE= 1821.854
VARIANCE= 1485.986
STANDARD DEVIATION= 38.548
MAXIMUM= 90.920
MINIMUM= -157.288

ABSOLUTE VALUE OF RESIDUALS
MEAN= 33.804
MEAN SQUARE= 1821.854
VARIANCE= 679.348
STANDARD DEVIATION= 26.064
MAXIMUM= 157.288
MINIMUM= 0.000

CHI-SQUARE= 80230.558